Norwegian Meteorological Institute
met.no
met.no report

no. 08/2004
Climate

Empirically downscaled SRES-based climate scenarios for Norway

RegClim results

Rasmus E. Benestad

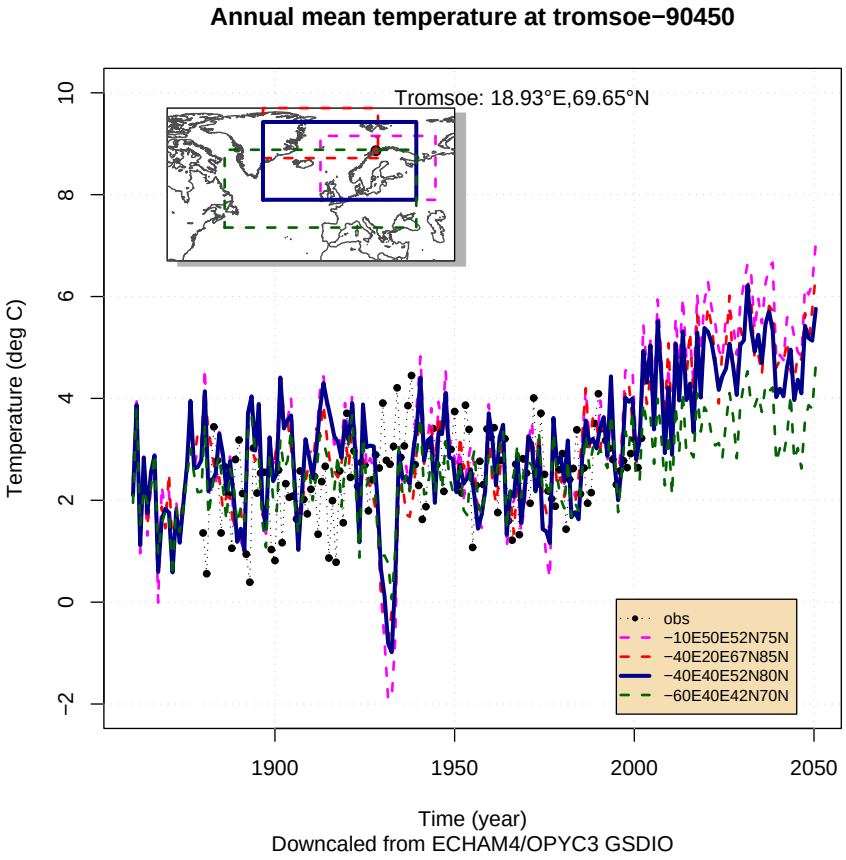

Figure 1: Example of empirical downscaling with the use of different domains

Norwegian Meteorological Institute
met.no

report

Title	Date
Empirically downscaled SRES-based climate scenarios for Norway	28th June 2004
Section	**Report no.**
Climate	no. 8/04
Author	**Classification**
R.E. Benestad	● Free ○ Restricted
	ISSN 1503-8025
	e-ISSN 1503-8025
Client(s)	**Client's reference**

Abstract

The results for 2000–2100 from six global climate models following the latest SRES B2 emission scenarios were examined in terms of common Empirical Orthogonal Functions (cEOFs). These common EOFs were also used as a basis for empirical downscaling, employing a stepwise multiple regression and a number of different predictors and domains. Statistics is presented for linear trend estimates and their associations with GCMs, predictors and domains. The sea level pressure (SLP) is not considered as an appropriated predictor for temperature, as the warming signal is not well captured by the SLP. The downscaling analysis indicated a general warming of the local climate, however, there were a few cases where negative temperature trends had been obtained. An explanation for these negative trends is that the predictor domains chosen were not appropriate.
The GCMs were generally not able to reproduce the observed annual cycle in the precipitation for interpolated locations. However, the downscaling analysis suggested generally good skill for the models using large-scale precipitation as predictor. Through empirical downscaling, local precipitation series with a realistic annual cycle can be constructed, but the large scatter in local seasonal precipitation variations interpolated from GCMs calls for the question whether current GCMs are able to predict how precipitation patterns will change under an enhanced greenhouse warming. There was no clear signal or consensus regarding future precipitation trends in traditional SLP based downscaling, but new downscaling models employing anomalous large-scale precipitation rates point to slight future trends in the precipitation for a selection of locations.

Keywords

Climate change, empirical downscaling, monthly mean temperature, monthly precipitation

Disciplinary signature	Responsible signature
Øystein Hov	Eirik Førland

Postal address P.O Box 43 Blindern, N-0313 OSLO, Norway
Office Niels Henrik Abels vei 40
Telephone +47 2296 3000
Telefax +47 2296 3050
e-mail: met.inst@met.no **Internet:** met.no
Bank account 7694 05 00601
Swift code DNBANOKK

Manuscript, met.no report 8/04, pp. 1–66

Title.

By R.E. Benestad

The Norwegian Meteorological Institute, PO Box 43, 0313, Oslo, Norway *

July 5, 2004

ABSTRACT

The results for 2000–2100 from six global climate models following the latest SRES B2 emission scenarios were examined in terms of common Empirical Orthogonal Functions (cEOFs). These common EOFs were also used as a basis for empirical downscaling, employing a stepwise multiple regression and a number of different predictors and domains. Statistics is presented for linear trend estimates and their associations with GCMs, predictors and domains. The sea level pressure (SLP) is not considered as an appropriated predictor for temperature, as the warming signal is not well captured by the SLP. The downscaling analysis indicated a general warming of the local climate, however, there were a few cases where negative temperature trends had been obtained. An explanation for these negative trends is that the predictor domains chosen were not appropriate.

The GCMs were generally not able to reproduce the observed annual cycle in the precipitation for interpolated locations. However, the downscaling analysis suggested generally good skill for the models using large-scale precipitation as predictor. Through empirical downscaling, local precipitation series with a realistic annual cycle can be constructed, but the large scatter in local seasonal precipitation variations interpolated from GCMs calls for the question whether current GCMs are able to predict how precipitation patterns will change under an enhanced greenhouse warming. There was no clear signal or consensus regarding future precipitation trends in traditional SLP based downscaling, but new downscaling models employing anomalous large-scale precipitation rates point to slight future trends in the precipitation for a selection of locations.

Key words: Climate change empirical downscaling monthly mean temperature monthly precipitation

*Corresponding author: R.E. Benestad, rasmus.benestad@met.no, The Norwegian Meteorological Institute, PO Box 43, 0313 Oslo, Norway, phone +47-22 96 31 70, fax +47-22 96 30 50

Contents

1 Introduction

Since the the downscaling of temperature and precipitation scenarios carried out by Benestad (2002a) a new set of emission-based climate scenarios have become available. The new set of emission scenarios are commonly referred to as the Special Report Emission Scenarios (SRES) (IPCC, 2001). This report updates the work by Benestad (2002a) with the new SRES-based climate scenarios.

2 Data & Methods

2.1 Data

The climate scenarios were downloaded from the IPCC Internet site* and subsequently converted to the netCDF format (Benestad, 2003c). The `R`-based statistical package `clim.pact` (Benestad, 2003a,b) was employed in order to compute EOFs and then to downscale the GCM results to local station values for the Nordic region. The empirical downscaling followed the procedure described in Benestad (2002a,b, 2001b), but using multiple regression instead of Canonical Correlation Analysis (CCA) for model calibration. The empirical downscaling was carried out for a set of stations (Tuomenvirta et al., 2001) in the Nordic countries, and the location of these stations are shown in Figure 1.

2.2 Method

The climatic trends presented in this report consist of linear best-fit evolution to time over the 2000-2100 interval, using the `R` regression model "`lm()`". The empirical downscaling was applied using adjusted EOFs (Imbert & Benestad, 2003) in order to correct for systematic GCM errors and to improve the match between the GCM simulations and the real world, unless otherwise stated. The EOF adjustment consisted of re-scaling the part of the PCs describing the GCM results so that they have the same mean and variance as the part describing the observations before year 2000.

Although Benestad (2003c) applied a "first-order" quality control to the SRES-based GCM results and found realistic description of the mean values, it is important to assess the realism of the models'

*http://ipcc-ddc.cru.uea.ac.uk/dkrz/dkrz_index.html

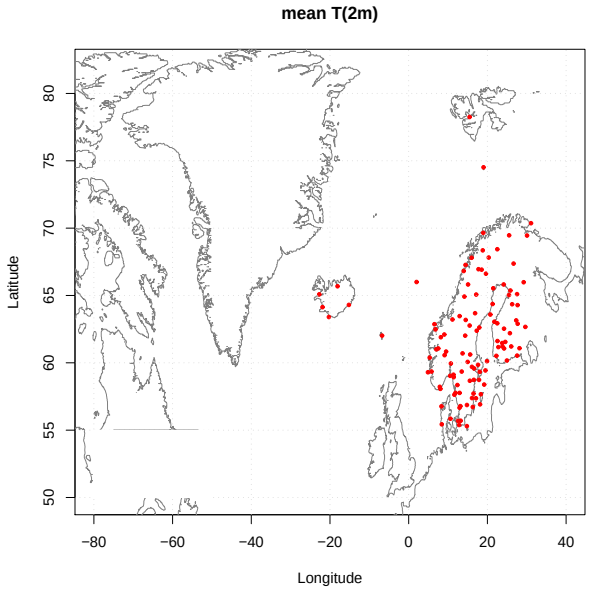

Figure 1. Map showing the location the Nordklim stations

variability and their description of the spatial structure of the climatic anomalies. By examining the common EOFs (Sengupta & Boyle, 1998; Barnett, 1999; Flury, 1988), a simple evaluation can be made which is directly relevant to the empirical downscaling, as the downscaling was carried out in the common EOF reference frame. Here, the evaluation was performed on mixed-common EOFs, which involve different entities as well as different data sources (i.e. observation and GCM). The 2-meter temperature [T(2m)] and sea level pressure (SLP) data were combined by synthesising new maps holding the maps of the T(2m) and SLP side-by-side in a similar fashion as the 'CPCA' approach described in Bretherton et al. (1992). The construction of these mixed fields was carried out using the `mixFields()`-function in the `clim.pact`-package for `R`. The mixed data sets were concatinated (along the time axis) in the same way as described by Sengupta & Boyle (1998); Barnett (1999); Flury (1988) (using the `catFields()`-function from `clim.pact`), and EOFs (henceforth referred to as 'mixed common EOFs') were estimated for these combined mixed fields (using the `EOF()`-function from `clim.pact`). Mixed-common EOFs has been employed before in empirical downscaling for local temperature on Svalbard (Benestad et al., 2002). The phrase 'common EOFs' (without 'mixed') will be used to denote EOFs estimated from a concatinated field (e.g. a gridded analysis and GCM results) but using one meteorological paramater only (e.g. SLP). The empirical downscaling was implemented with the `DS()`-function from `clim.pact` (e.g. see the examples given in the documentation of the `clim.pact`-package for `R`).

3 Results

3.1 Quality control: common EOFs

Figures 2 – 7 show January unadjusted mixed-common EOF diagnostics for the ECHAM4/OPYC3 GSDIO integration, the ECHAM4/OPYC3 B2 SRES integration, the CCCma B2 SRES integration, the CSIRO B2 SRES integration, the HadCM3 B2 SRES integration, and the NCAR-PCM B2 SRES integration respectively. A superficial inspection suggests similar large-scale feature in the ECHAM4/OPYC3 and the CCCma results. However, the part of the PCs representing the CCCma results has smaller amplitude than the observations (Figure 4b). The CSIRO GCM appears to give less realistic temperature pattern and smoother SLP anomalies (Figure 5a). These spatial structures are dominated by the GCM results because the GCM record was longer than that of the observations. The HadCM3 results give more prominent small-scale details in the temperature field than the other models, and SLP field hints to a realistic storm track (Figure 6a). The NCAR-PCM results suggest a presence of systematic errors in the SLP anomalies, and problems describing the land-sea temperature contrasts (Figure 7a). The strong weights over ocean areas in the temperature pattern may be a result of unrealistically cold conditions in the models description of the present-day climate and a southward bias in the ice-edge location. These EOF structures are dominated by the GCM because of the simulated longer record, and the amplitude of the leading PC weights are only slightly greater for the GCM than for the observations (Figure 7b).

Panel c in Figures 2 – 7 show show scatter plots between the leading two PCs. The two leading PCs exhibit similar distributions for all the models, except for CCCma where the weights in the leading PC are weaker for the GCM than the NCEP re-analysis. For the second PC, the distribution of the NCEP values exhibits three outliers (Figure 4c). Panel d in Figures 2 – 7 show show the proportional variance (eigenvalues) associated with the leading EOFs. The two leading modes were well separated (North et al., 1982) for the NCEP-ECHAM4/OPYC3 and the NCEP-HadCM3 EOFs, but not for the other GCMs. Although the first and second EOFs were not resolved for the NCEP-NCAR-PCM data according to North et al. (1982), their error bars barely overlapped. The second and third modes were clearly degenerate (a 'pair' with similar values) in the various mixed-common EOF results except for the CSIRO model.

In summary, the ECHAM4/OPYC3 and HadCM3 models appear to give the most realistic description of the anomalies over northern Europe.

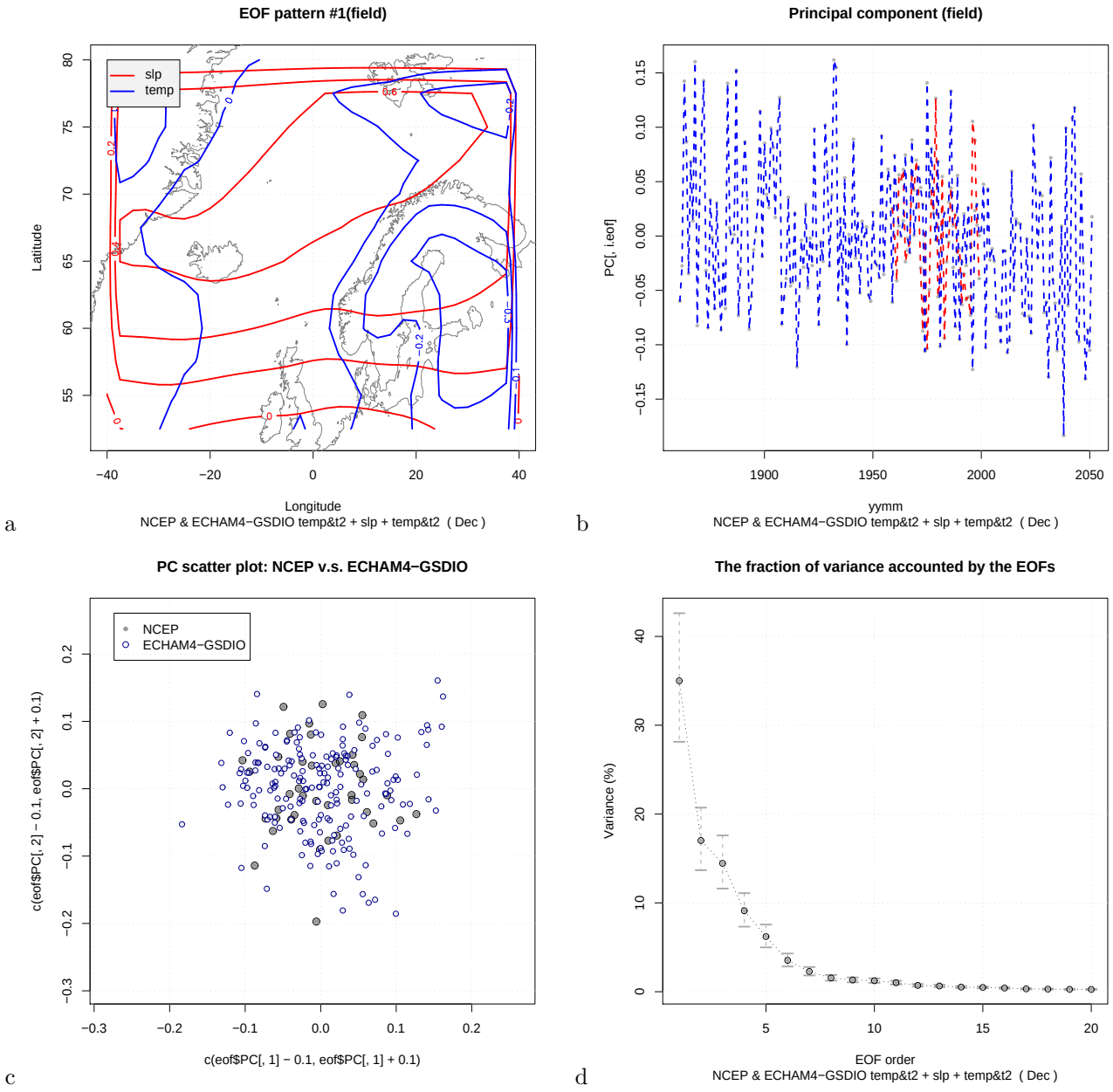

Figure 2. Mixed-common EOF products from the ECHAM4/OPYC3 GSDIO integration for January. (a) leading spatial EOF patterns, (b) leading PC (red= NCEP; blue= GCM), (c) comparison between scatter plots between the two leading PCs from NCEP reanalysis (1958–1998) and the GCM ("1860"–"2050"), (d) the proportional variance described by the leading modes.

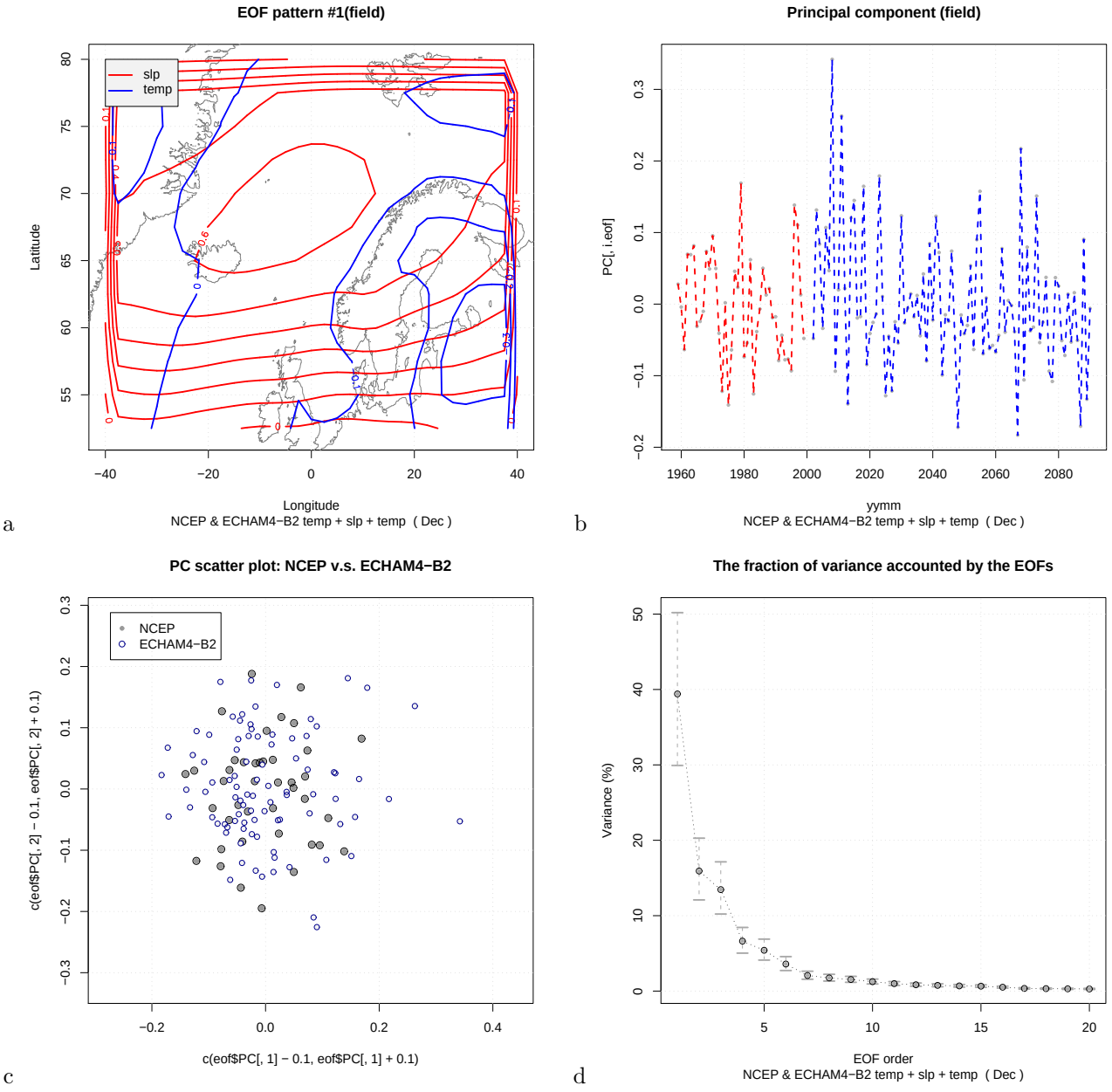

Figure 3. Mixed-common EOF products from the ECHAM4/OPYC3 B2 SRES integration for January. (a) leading spatial EOF patterns, (b) leading PC (red= NCEP; blue= GCM), (c) comparison between scatter plots between the two leading PCs from NCEP reanalysis (1958–1998) and the GCM (see Table 4 for time intervals), (d) the proportional variance described by the leading modes.

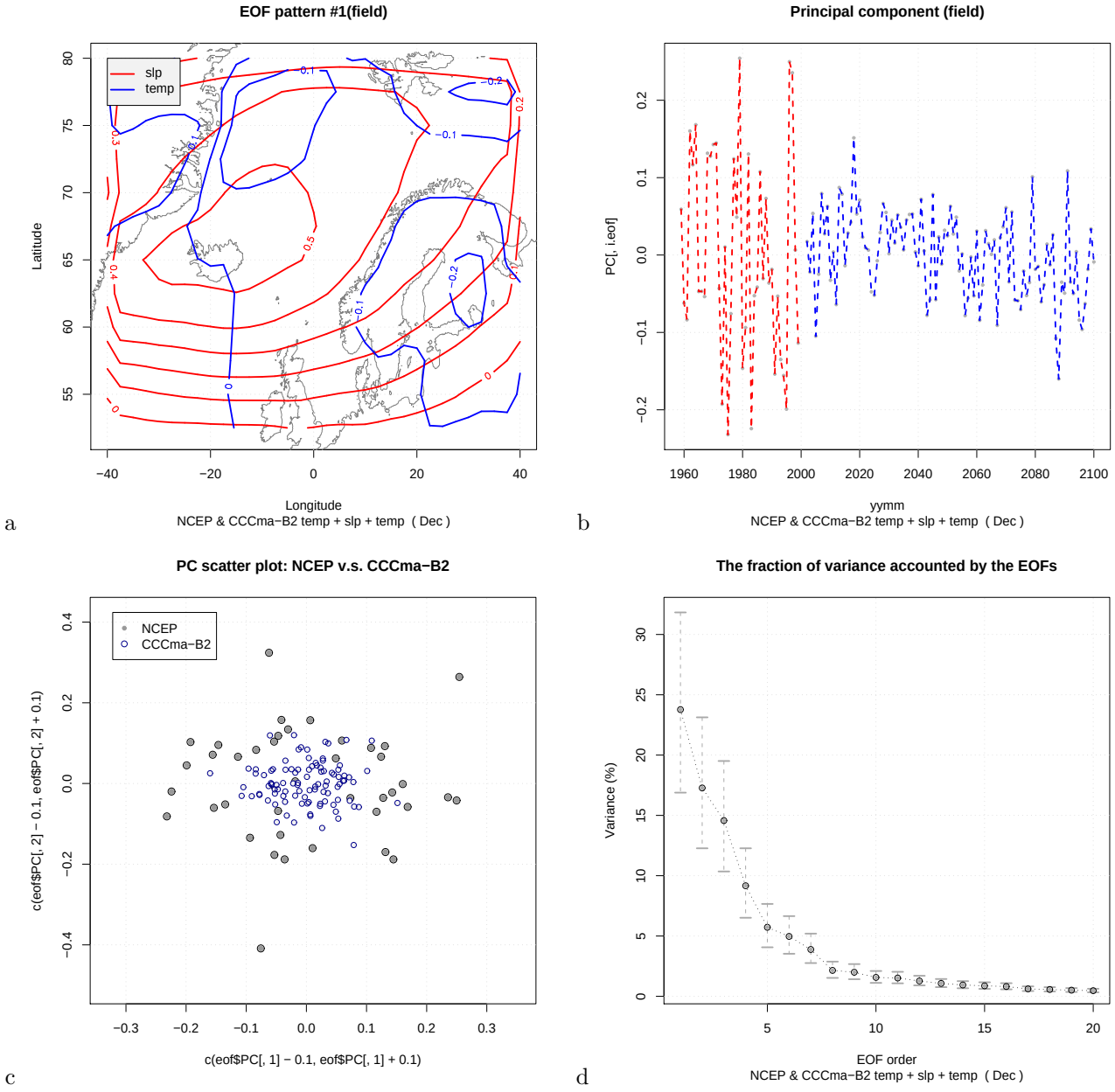

Figure 4. Mixed-common EOF products from the CCCma B2 SRES integration for January. (a) leading spatial EOF patterns, (b) leading PC (red= NCEP; blue= GCM), (c) comparison between scatter plots between the two leading PCs from NCEP reanalysis (1958–1998) and the GCM (see Table 4 for time intervals), (d) the proportional variance described by the leading modes.

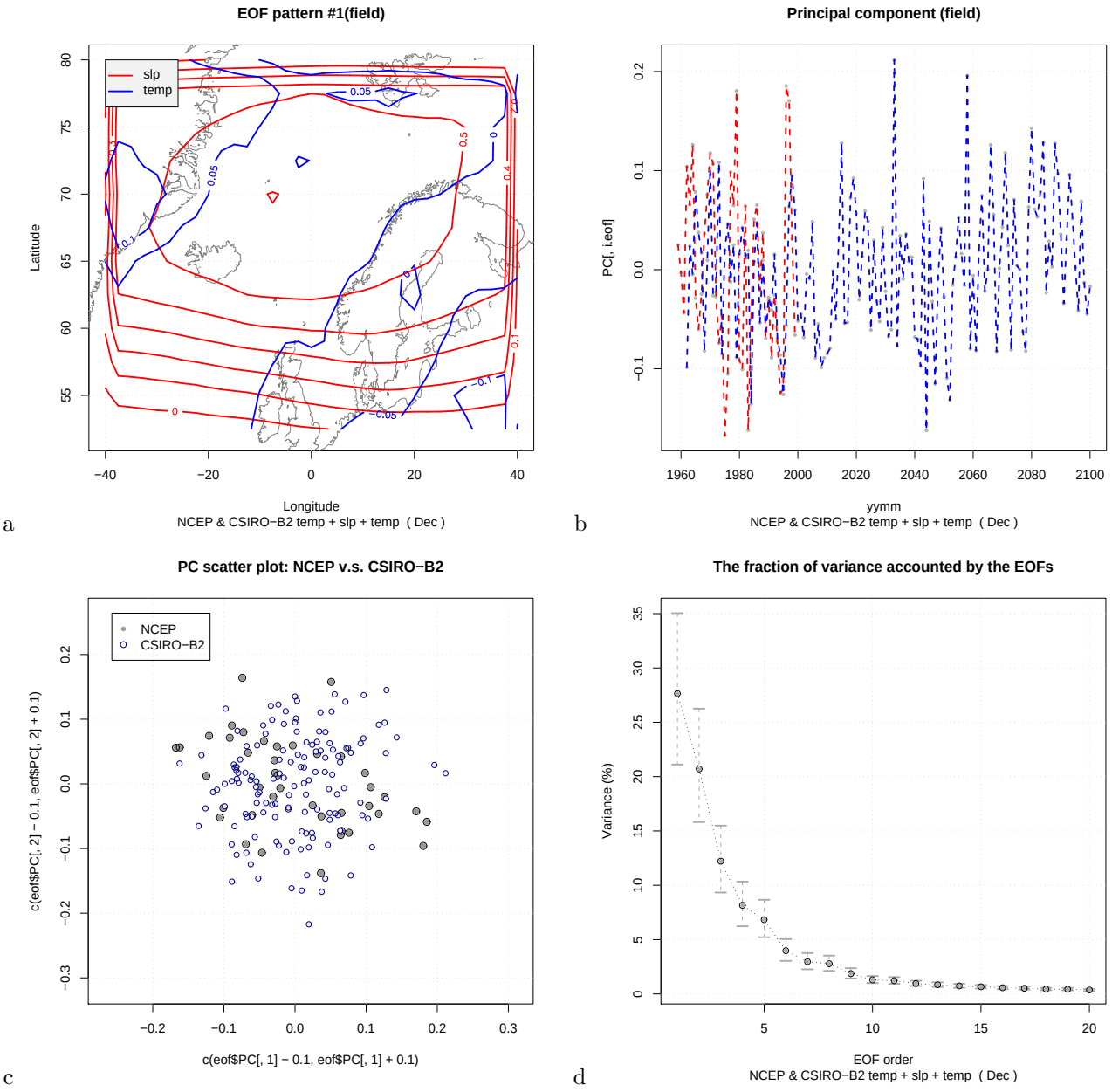

Figure 5. Mixed-common EOF products from the CSIRO B2 SRES integration for January. (a) leading spatial EOF patterns, (b) leading PC (red= NCEP; blue= GCM), (c) comparison between scatter plots between the two leading PCs from NCEP reanalysis (1958–1998) and the GCM (see Table 4 for time intervals), (d) the proportional variance described by the leading modes.

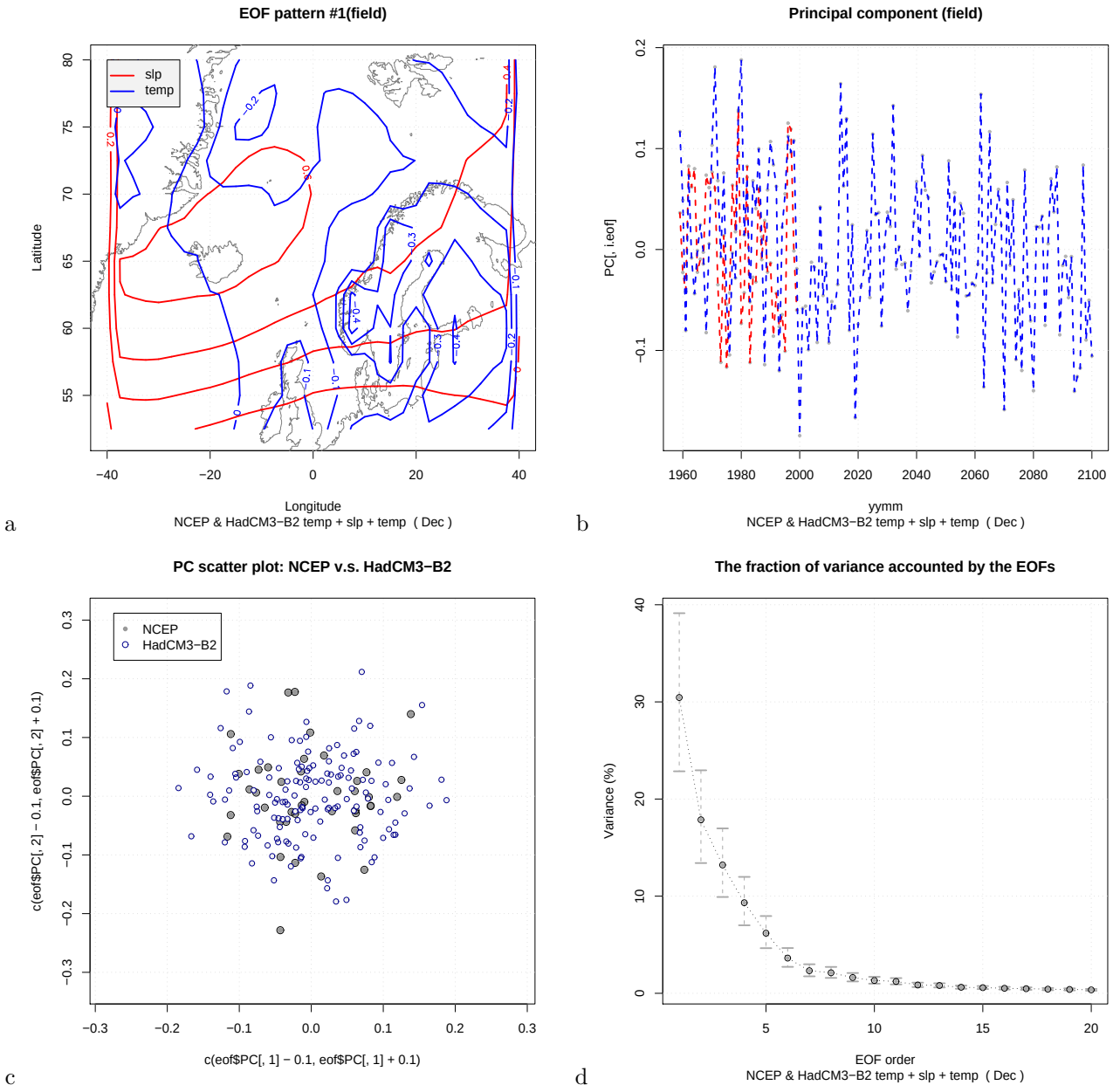

Figure 6. Mixed-common EOF products from the HadCM3 B2 SRES integration for January. (a) leading spatial EOF patterns, (b) leading PC (red= NCEP; blue= GCM), (c) comparison between scatter plots between the two leading PCs from NCEP reanalysis (1958–1998) and the GCM (see Table 4 for time intervals), (d) the proportional variance described by the leading modes.

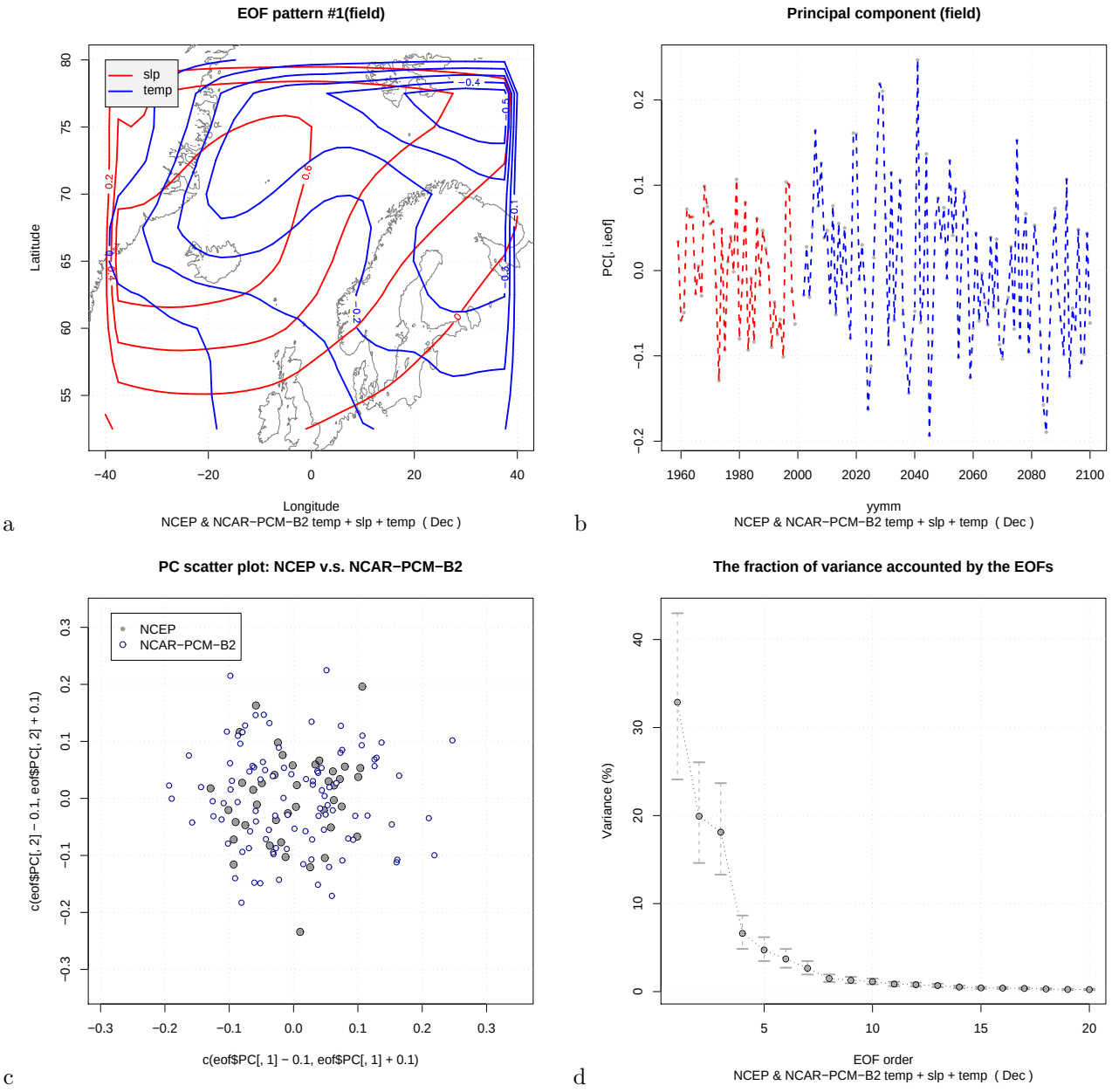

Figure 7. Mixed-common EOF products from the NCAR-PCM B2 SRES integration for January. (a) leading spatial EOF patterns, (b) leading PC (red= NCEP; blue= GCM), (c) comparison between scatter plots between the two leading PCs from NCEP reanalysis (1958–1998) and the GCM (see Table 4 for time intervals), (d) the proportional variance described by the leading modes.

3.2 Skill & trend statistics

Summary statistics were produced for the downscaled results of the climate model output following the SRES B2 emission scenario in order to get a quick overview over the main downscaled characteristics. The predictands in this study were taken from the Nordklim data (Tuomenvirta et al., 2001) (see Figure 1). Figure 8 shows a measure of the predictor skill associated with the various predictor choices, here the R^2-estimates associated with the multiple regression between the local series and the large-scale anomalies. The R^2 values indicate how much of the variance the downscaled results describe, where 0% denotes no skill and 100% means a perfect reproduction. Panel (a) shows the R^2 estimates for the temperature for all calendar months (January–December) and panel (b) shows the same analysis for the precipitation. anels (c) and (b) show the downscaled linear trend estimates associated with the various predictors for temperature and precipitation respectively. The interpretations of Figure 8(a) are that temperature-based (purely temperature or mixed-field predictor including temperature) are associated with the best description of the link between the large and small temperature scales. Precipitation rate and SLP give similar skill in the region $R^2 \sim$ 40–70% in describing the link between large and small scales, whereas humidity-based predictors yield lower scores. Panels (c) and (d) show the trend statistics related to the various predictors. SLP gives significantly lower trend estimates than the temperature-based predictors. The various predictors for precipitation do not distinguish much from one another, with the exception of the mixed field giving a greater spread.

Figure 9 shows trend statistics related to the various GCMs, each box-and-whisker entity based on all the Nordklim stations. The GCM-based differences for indivudual sites will be resolved later on in this paper. All the climate models indicate a warming of a similar magnitude. Hence, there is a good agreement amongst the various GCMs regarding the general warming. The same statistics for the downscaled precipitation trends (panel b), shows that all the GCMs indicate a scatter centred around zero, but with different range for different GCMs. CSIRO, ECHAM4/OPYC3 and GFDL suggest a greater range (stronger trends) than CCCma, HadCM3 and NCAR-PCM. Panels (c) and (d) show trend statistics associated with different predictor domain choices. The multi-model ensemble results (all Nordklim stations) do not appear, from these statistics, to be very sensitive to the predictor domain.

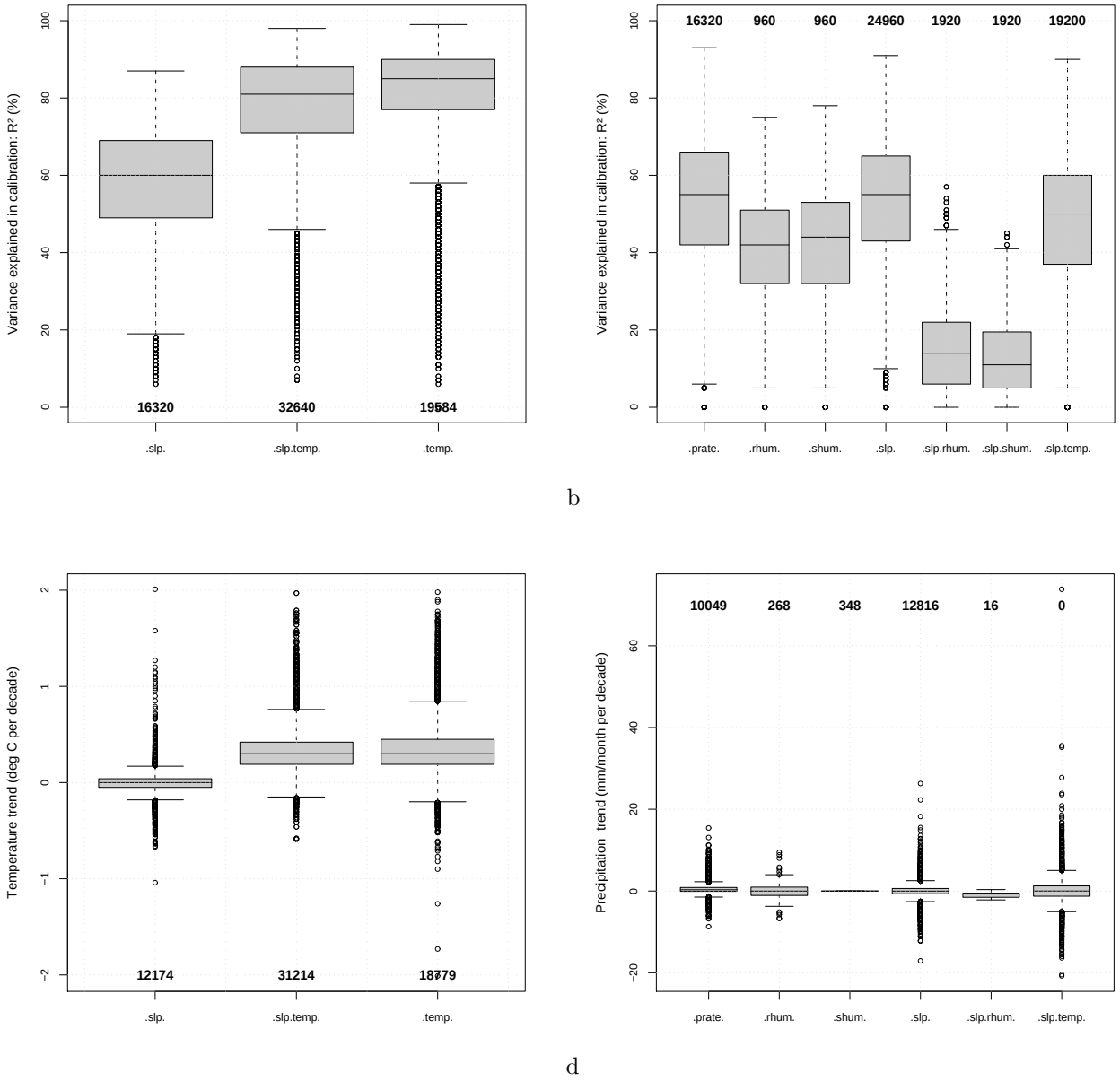

Figure 8. Box-whisker plots of R^2-estimates dervied from the multiple regresion between the local series and large-scale anomalies (a,b), and linear trend rates (in time, the unit is °C/decade or mm/month per decade) associated with varous predictors. The box shows the interquantile range and the whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. The outliers are shonw seperately.

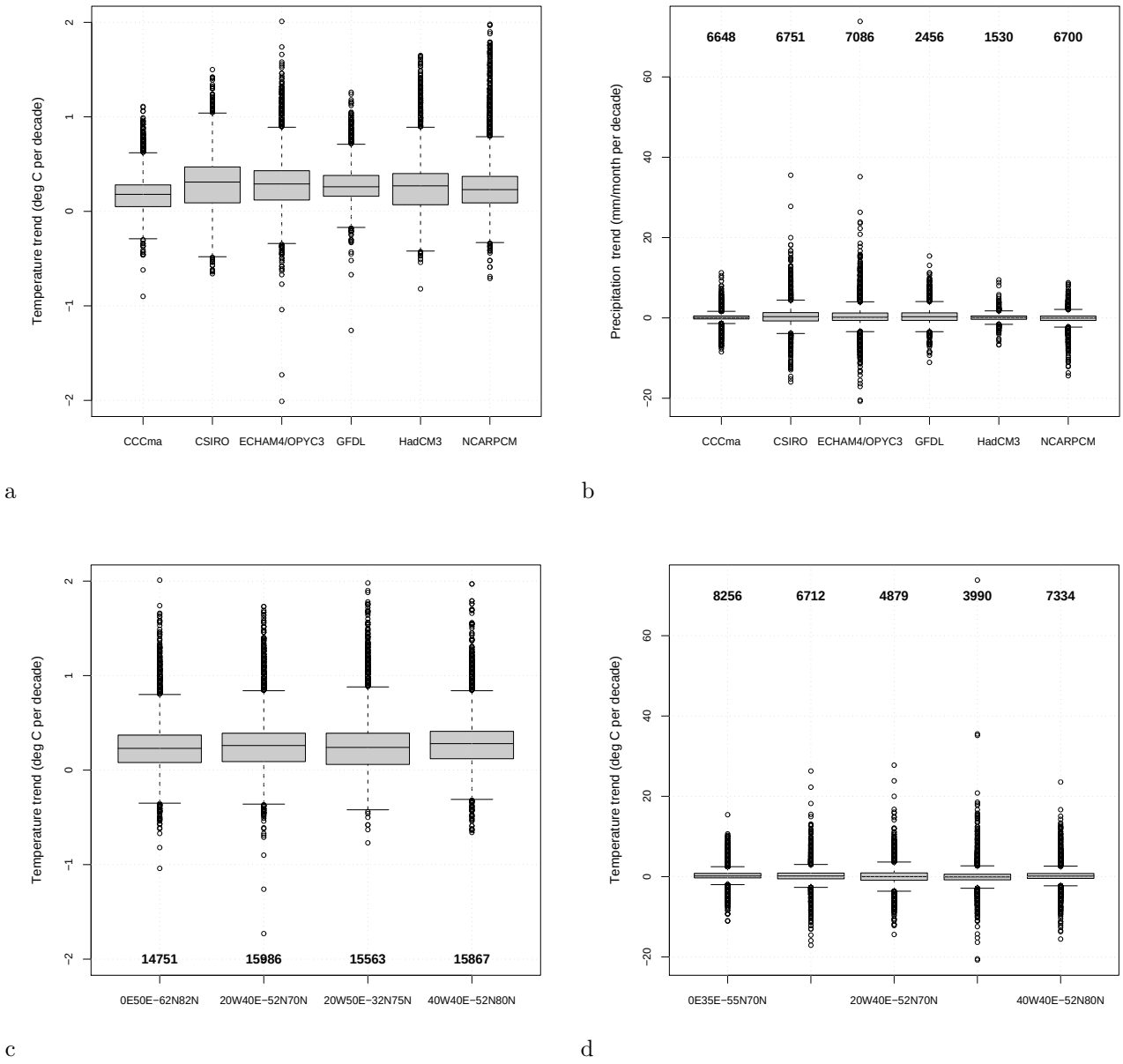

Figure 9. Box-whisker plots whowing linear trend rates (in time) associated with the individual GCMs (a,b) and different predictor domain (c,d). The unit of the tend rates are °C/decade and mm/month per decade for panels (a,c) and (b,d) respectively. The box shows the interquantile range and the whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. The outliers are shown seperately.

3.3 Cases with a projected cooling

Although the statistics indicate a general warming, there are some cases where negative temperature trends have been calculated for the downscaled SRES B2 scenarios. It is important to look in more detail on the cases where a cooling is obtained in climate scenarios that are known to describe a warming trend. It is expected that predictors based solely on SLP, which does not capture the general warming signal well (Figure 8(c)), may result in a number of cases with negative temperature trends due to random variability (year-to-year variations). Table 1 lists the number of negative trends associated with the different predictors, and the SLP-only predictor is responsible for most of these cases. However, there are a number times where the temperature-based predictors also produce a negative temperature trend, contrary to expectations. In order to examine some of these, the next summary of statistics is limited to the temperature-only predictor and to the cases with R^2 >50%. Tables 2 and 3 show the number of negative temperature trends associated with the different GCM, domain, and calendar month. Most of the negative trends (30%) are associated with the Canadian climate model CCCma whereas the German ECHAM4/OPYC3 model accounted for 8%. The CCCma model had lower spatial resolution than the others, with the exception of the Australian CSIRO model (Table 4). There is also a tendency for the negative trend estimates being more common during winter when the inter-annual variability is most pronounced (Table 3).

Table 5 shows the number of negative trend-estimates according to location. The locations with most cases included Bjørnøya (36), Svalbard (32), Akureyri (29), Vardø (28), Jan Mayen, Karasjok, (26), Tromsø (25), Nesbyen (24), Vestervig (24), and Stykkisholmur (22). The predictor domains were either 20°W–40°E & 52°N–70°N or 40°W–40°E & 52°N–80°N. The former domain does not extend sufficiently far north as to make it an appropriate choice for empirical downscaling of Arctic stations. Figure 10 illustrates how domains that do not sufficiently cover the vicinity of the stations may produce too low warming trends. Benestad (2002b) warned against the danger in applying a 'blind' analysis without inspecting the predictor patterns e.g. for physical plausibility, and illustrated how cooling trends can be inferred for Greenlandic stations when the predictor domain focuses on northern Europe. In that case, the downscaling models (correctly) identifies the North Atlantic Oscillation (NAO) pattern with an anti-correlation between the temperature in western Greenland and northern Europe as the main feature. If the GCMs project a stronger warming over northern Fennoscandia for the future, then the established empirical relationship can produce a misleading picture with a (spurious) cooling over Greenland. Since it is practically impossible to manually check the plausibility of the large sample of downscaled trends presented here, this analysis admitedly suffers from being 'blindly' downscaled. The predictor domains are nevertheless believed to be appropriate for a large number of the stations presented here if the Arctic stations are excluded (a different set of predictor domains should be used for the Arctic stations).

		.slp.	.slp.temp.	.temp.	total
CCCma	count	795	298	235	13192
CSIRO	count	1460	140	81	13192
ECHAM4/OPYC3	count	1139	166	64	13192
GFDL	count			126	3332
HadCM3	count	1209	160	159	13056
NCARPCM	count	1096	94	124	13056
all	count	5699	858	789	69020
all	proportion	35%	3%	4%	100%

TABLE 1. Number of cases with a negative temperature trend for each predictor type. Here all the trend estimates for the SRES A2 (476) and B2 (68544) scenarios were included. The total number of trend estimates is 170544, of which 24604 (14.4%) were negative. The time interval over which the linear trend was estimated was the same as in Figure 3–7 (see Table 4 for time intervals).

month	CCCma	CSIRO	ECHAM4/OPYC3	GFDL	HadCM3	NCARPCM
Jan	43	4	8	19	17	32
Feb	44	3	5	12	78	11
Mar	16	6	3	5	7	23
Apr	9	1	0	9	0	4
May	18	12	3	11	4	4
Jun	8	7	16	13	5	3
Jul	25	10	1	10	0	9
Aug	11	2	2	2	1	3
Sep	8	10	7	4	7	1
Oct	8	9	8	9	2	19
Nov	22	7	6	23	14	3
Dec	21	8	5	7	23	10
all	233	79	64	124	158	122

TABLE 2. Number of cases with a negative temperature trend for each GCM. Here only the trends estimates derived using the temperature only predictor (".temp.") and with R^2 >50% were included (780 [4.0%] cases of negative trend in total of in the category ".temp." and R^2 >50%). SRES A2 (476) and B2 (68544).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
D1	70	77	32	15	6	3	41	8	8	8	30	22
D2	24	29	20	8	20	8	11	12	10	39	22	27
D3	9	29	6	0	20	32	0	1	16	4	20	19
D4	20	18	2	0	6	9	3	0	3	4	3	6
all	123	153	60	23	52	52	55	21	37	55	75	74

TABLE 3. Number of cases with a negative temperature trend for each calendar month. Here only the trends estimates derived using the temperature only predictor (".temp.") and with R^2 >50% were included. The domains are definedd as follows: D1=0E50E–62N82N, D2=20W40E–52N70N, D3=20W50E–32N75N, and D4= 40W40E–52N80N. SRES A2 (476) and B2 (68544). Total number of data is 69020, which includes 68 stations.

Model	Flux adj.	ny×nx	country	A2	B2	period
CCCma	Yes	96 × 48	Canada	136	13056	1900–2100
CSIRO Mk2	Yes	64 × 56	Australia	136	13056	1961–210
ECHAM4/OPYC3	Yes	128 × 64	Germany	136	13056	1990–2100
GFDL-30	No	96 × 80	U.S.A.	68	3264	
GFDL-30 (SLP)		192 × 80				
HadCM3	No	96 × 73	U.K.	0	13056	1950–2099
NCAR-CSM	No	128 × 64	U.S.A.	0	0	2000–2099
NCAR-PCM	No	128 × 64	U.S.A.	0	13056	1980–2099

TABLE 4. A summary of the GCMs and their spatial resolution. The notations nx and ny denote the number of grid boxes along the zonal and meridional directions respectively.

	ABISKO	AKUREYRI	BERGEN-FLORIDA	BJOERNOEYA
D1	3	10	4	4
D2	9	2	0	27
D3	2	15	2	4
D4	0	2	0	1
all	14	29	6	36
	BORAAS	FALSTERBO	FALUN	FERDER
D1	6	9	4	6
D2	0	2	0	4
D3	2	1	0	4
D4	1	2	0	0
all	9	14	4	14
	GLOMFJORD	GOETEBORG	GOTSKA	HAERNOESAND
D1	0	8	4	2
D2	1	0	0	0
D3	2	2	0	2
D4	0	2	0	0
all	3	12	4	4
	HALMSTAD	HAMMERODDE	HAPARANDA	HELSINKI
D1	9	10	2	4
D2	1	2	2	1
D3	1	1	1	1
D4	2	4	0	0
all	13	17	5	6
	HOBURG	HOLMOEGADD	JAN MAYEN	JOKKMOKK
D1	5	4	5	2
D2	0	0	20	1
D3	1	1	0	1
D4	0	1	1	0
all	6	6	26	4
	JYVAESKYLAE	KAJAANI	KALMAR	KARASJOK
D1	2	2	7	2
D2	3	1	0	20
D3	1	2	1	4
D4	1	3	3	0
all	7	8	11	26
	KARESUANDO	KARLSTAD	KJOEREMSGRENDI	KOEBENHAVN
D1	3	2	9	10
D2	12	1	1	1
D3	4	1	3	2
D4	0	0	0	2
all	19	4	13	15
	KUOPIO	KUUSAMO	KVIKKJOKK	LAERDAL
D1	2	0	3	7
D2	2	6	4	3
D3	0	1	1	1
D4	2	1	0	2
all	6	8	8	13
	LANDSORT	LAPPEENRANTA	NESBYEN	NORDBY
D1	5	2	13	6
D2	0	5	3	1
D3	0	0	6	2
D4	0	2	2	2
all	5	9	24	11
	OELANDS	OESTERSUND	OKSOEY	ONA
D1	8	3	7	3
D2	1	2	2	1
D3	0	1	1	4
D4	3	0	1	1
all	12	6	11	9
	OSLO-BLINDERN	OULU	PITEAA	REYKJAVIK
D1	5	0	3	6
D2	1	2	0	2
D3	1	1	1	7
D4	0	1	1	4
all	7	4	5	19
	SHIP M	SODANKYLAE	STENSELE	STOCKHOLM
D1	2	1	0	7
D2	0	4	0	1
D3	2	2	0	1
D4	1	0	0	0
all	5	7	0	9

TABLE 5. Number of cases with a negative temperature trend (see Table 4 for intervals) for each calendar location. Here only the trends estimates derived using the temperature only predictor (".temp.") and with R^2 >50% were included (780 cases of negative trend in the grand total of 62,167, i.e. 1%). The domains are defined as follows: D1=0E50E–62N82N ($N = 17612$), D2=20W40E–52N70N ($N = 17136$), D3=20W50E–32N75N (N=17136), and D4= 40W40E–52N80N ($N = 17136$). SRES A2 (476) and B2 (68544) and from 6 GCMs.

	STYKKISHOLMUR	SVALBARD	SVEG	SVENKSA
D1	7	1	5	7
D2	2	21	1	1
D3	11	9	2	0
D4	2	1	0	1
all	22	32	8	9
	TAERNABY/HEMAVAN	TAMPERE	TEIGARHORN	TORSHAVN
D1	1	2	4	4
D2	0	1	0	0
D3	1	1	6	7
D4	0	0	1	1
all	2	4	11	12
	TRANEBJERG	TROMSOE	TURKU	UPPSALA
D1	10	0	4	5
D2	1	18	1	0
D3	1	6	0	0
D4	2	1	1	1
all	14	25	6	6
	UTSIRA	VAERNES/TRONDHEIM	VAEXJOE	VARDOE
D1	7	2	7	0
D2	1	3	0	22
D3	2	3	1	3
D4	2	1	2	3
all	12	9	10	28
	VESTERVIG	VESTMANNAEYAR	VINGA	VISBY
D1	13	4	7	9
D2	3	3	1	0
D3	5	4	0	1
D4	3	2	2	0
all	24	13	10	10

TABLE 5. Number of cases with a negative temperature trend for each calendar location. Here only the trends estimates derived using the temperature only predictor (".temp.") and with R^2 >50% were included (780 cases of negative trend in the grand total of 62,167, i.e. 1%). The domains are defined as follows: D1=0E50E–62N82N ($N = 17612$), D2=20W40E–52N70N ($N = 17136$), D3=20W50E–32N75N (N=17136), and D4= 40W40E–52N80N ($N = 17136$). SRES A2 (476) and B2 (68544) and from 6 GCMs.

	0E50E–62N82N	20W40E–52N70N	20W50E–32N75N	40W40E–52N80N	total
CCCma	78	75	36	44	
CSIRO	9	49	13	8	
ECHAM4/OPYC3	12	19	31	2	
GFDL	70	21	24	9	
HadCM3	71	34	48	5	
NCARPCM	80	32	4	6	
all	320	230	156	74	

TABLE 6. Number of cases with a negative temperature trend for each calendar month. Here only the trends estimates derived using the temperature only predictor (".temp.") and with R^2 >50% were included. SRES A2 (476) and B2 (68544) and from 6 GCMs.

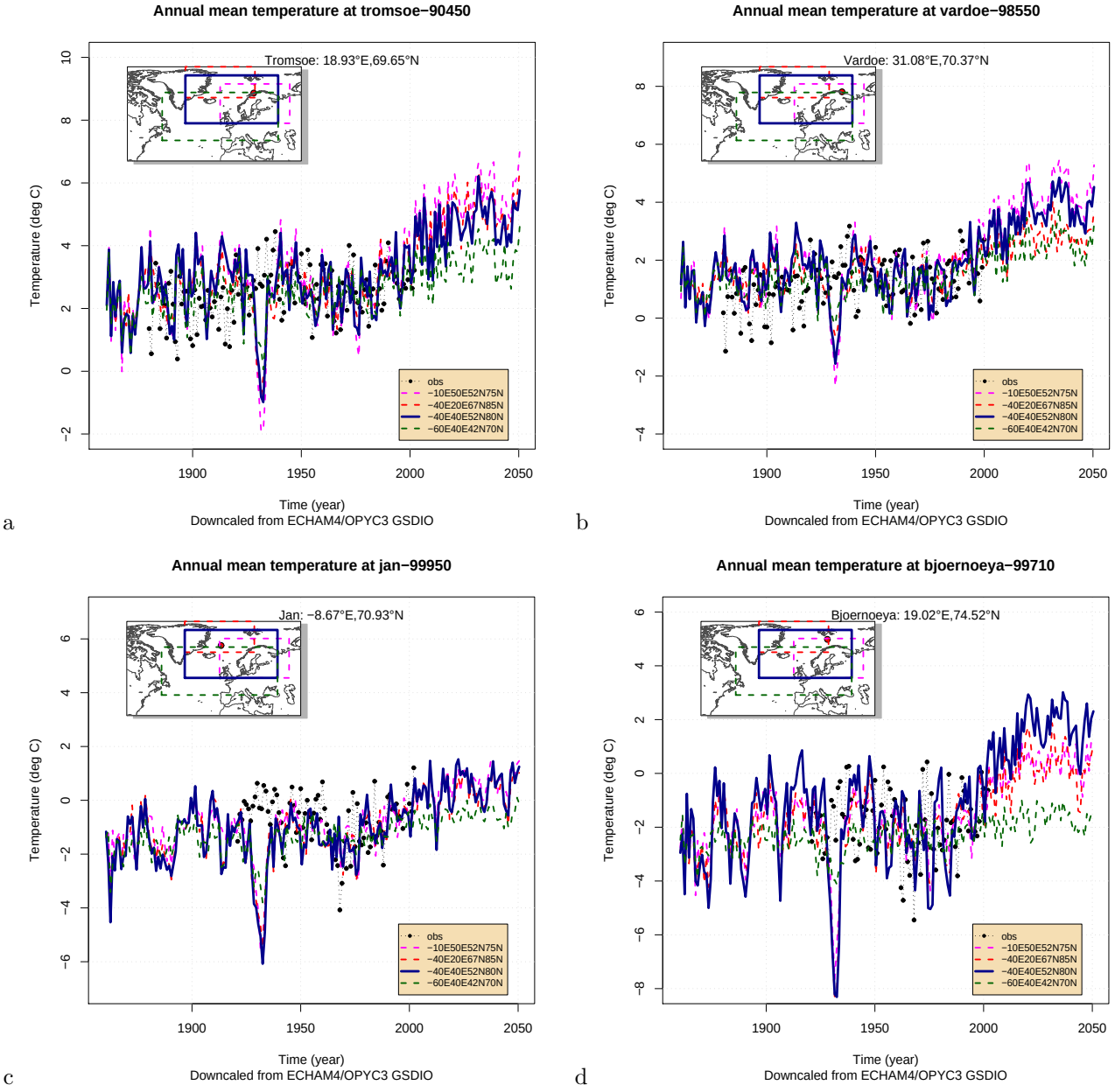

Figure 10. Downscaled annual mean temperature at Tromsø (a), Vardø (b), Jan Mayen (c), and Bjørnøya (d) using different domains. These scenarios were derived from the ECHAM4/OPYC3 GSDIO integration following the IS92a emission scenario (as opposed to the SRES B2 emission scenario for the rest of the results).

3.4 Downscaled temperature

Figure 11 shows seasonally stratified monthly mean temperatures downscaled for Oslo-Blindern. In contrast to Figure 9(a), the different GCMs indicate slightly different temperature estimates for the individual stations and seasons. The CCCma model has a tendency to produce estimates in the lower part of the range from all GCMs. Also contrary to the impression from earlier studies based on IS92a emission scenarios (Benestad, 2002b, 2004, 2002a, 2000b, 2001a), the winter-time warming does not appear to be the strongest: here the magnitude of the spring-time warming is similar, apparently in closer agreement with the observed trends in Norway (Benestad, 2001a). The GFDL model gives the impression of smallest scatter, but this is because only the '.temp.' predictor was used for downscaling this GCM (SLP was not available as only the surface pressure was available).

Figure 12 and 13 show the time series of winter temperature downscaled for Oslo, Tromsø, Bergen, Reykjavik (Iceland), Helsinki (Finland), Jokkmokk (Sweden), Stockholm (Sweden), and Copenhagen (Denmark), based on the different GCMs. Most of the scenarios indicate realistic variance, apart from the CCCma which has a tendency to produce too weak inter-annual variations.

Figure 14 shows the same analysis as in Figure 11a, but for Tromsø, Reykjavik, Stockholm and Abisko instead of Oslo. The estimates derived from the NCAR-PCM model tend to have a slightly greater scatter than for the other models. The different GCMs tend to indicate similar future winter warming in Tromsø, albeit with slightly weaker values for CCCma. At Stykkisholmur, there are some variations amongst the different GCM-based estimates, with strongest warming in NCAR-PCM and GFDL. The estimates based on the GFDL GCM have a smaller range than the others for Abisko, due to the exclusion of SLP and the SLP-T(2m) mixed fields in the GFDL predictor set.

Figure 15– 20 present the results from Figures 11 & 14 as histograms and provides a crude measure of probabilities associated with the warming at Oslo-Blindern, Tromsø, Bergen-Florida, Helsinki, Abisko and Copenhagen, given the SRES B2 emission scenario and assuming the climate models and the empirical downscaling are unbiased. These histograms show the distribution of all the trend estimates, also those derived using only SLP as predictor. There is a bias towards zero trend in the SLP-based results, which does not reflect the underlying CO_2-related warming trend (Table 1 and Figure 8(c)). The tail of the distribution has a tendency to protude further into negative values in winter than for the other seasons, in agreement with the findings from Table 3. There are hints of secondary peaks around zero in the distributions of winter, spring and summer temperatures, however, this feature is less pronounced in autumn. The secondary peak around zero due to the SLP-only predictor is present in many of these plots and therefore these distributions are biased. Hence, the actual warming projected by the models is expected to be higher than that deduced from these histograms. The SLP-only predictor is therefore not an appropriate predictor for local temperature.

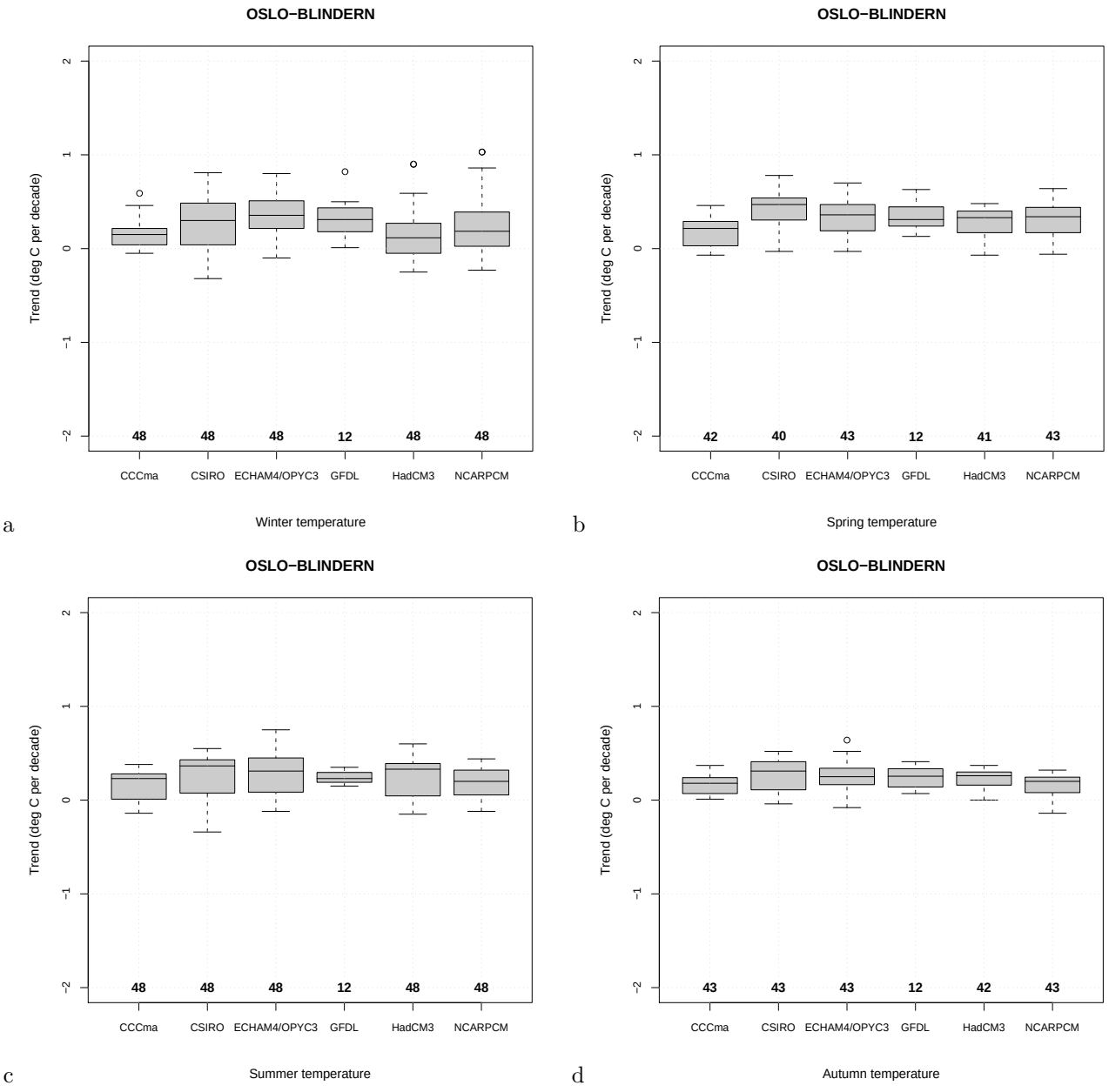

Figure 11. Downscaled monthly mean temperature trends for Oslo-Blindern for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The number of data points in each box is denoted near the bottom axis.

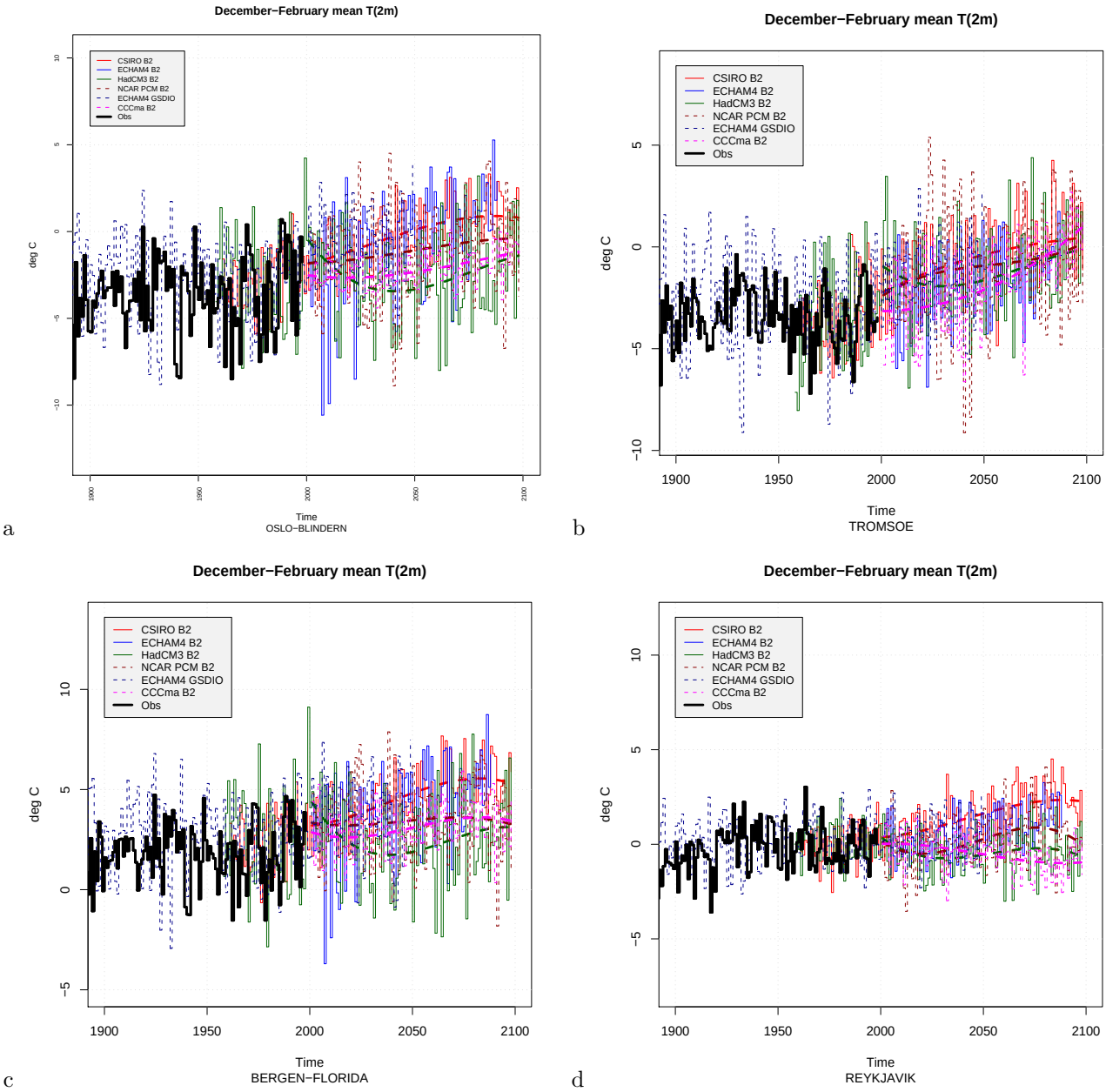

Figure 12. Downscaled winter temperature trends for Oslo-Blindern (a), Tromsø (b), Bergen-Florida (c) and Reykjavik (d). The trends estimates derived using all predictor types and domains, but only R^2 >50% were included.

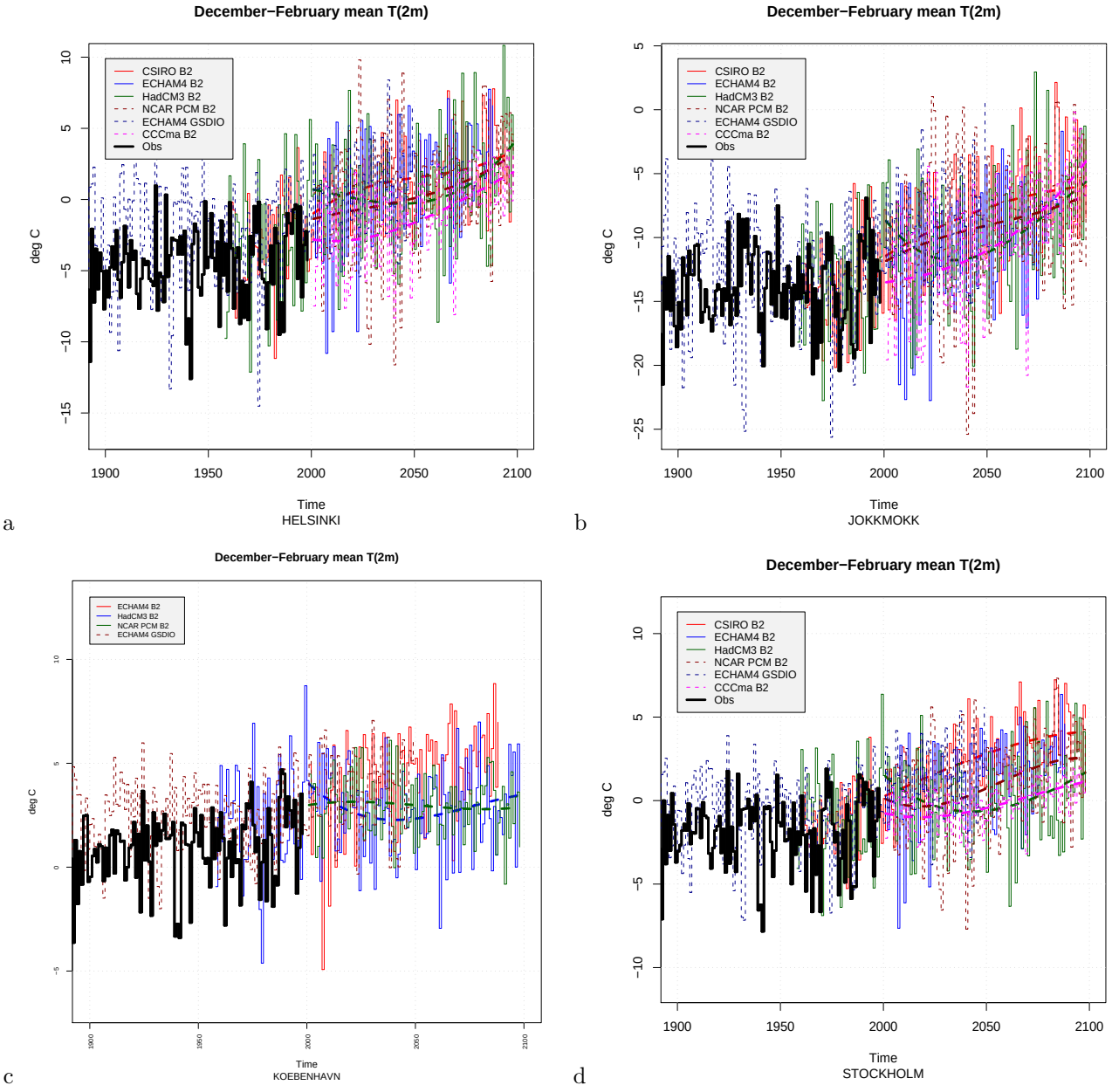

Figure 13. The same as Figure 12, but for Helsinki (a), Jokkmokk (b), Copenhagen (c) and Stockholm (d).

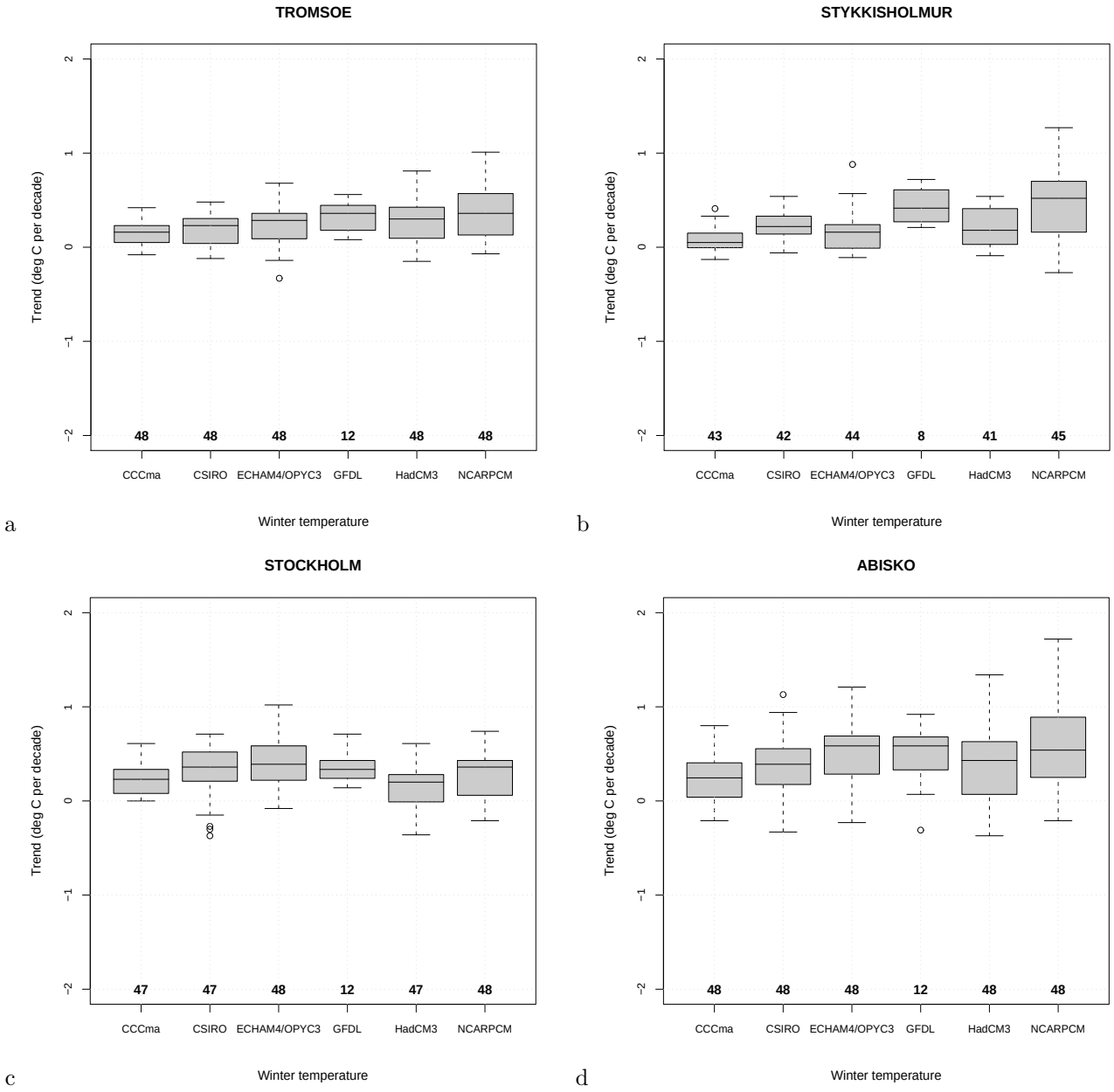

Figure 14. Downscaled winter temperature trends similar to Figure 11(a), but for for Tromsø (a), Stykkisholmur (b), Stockholm (c) and Abisko (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included.

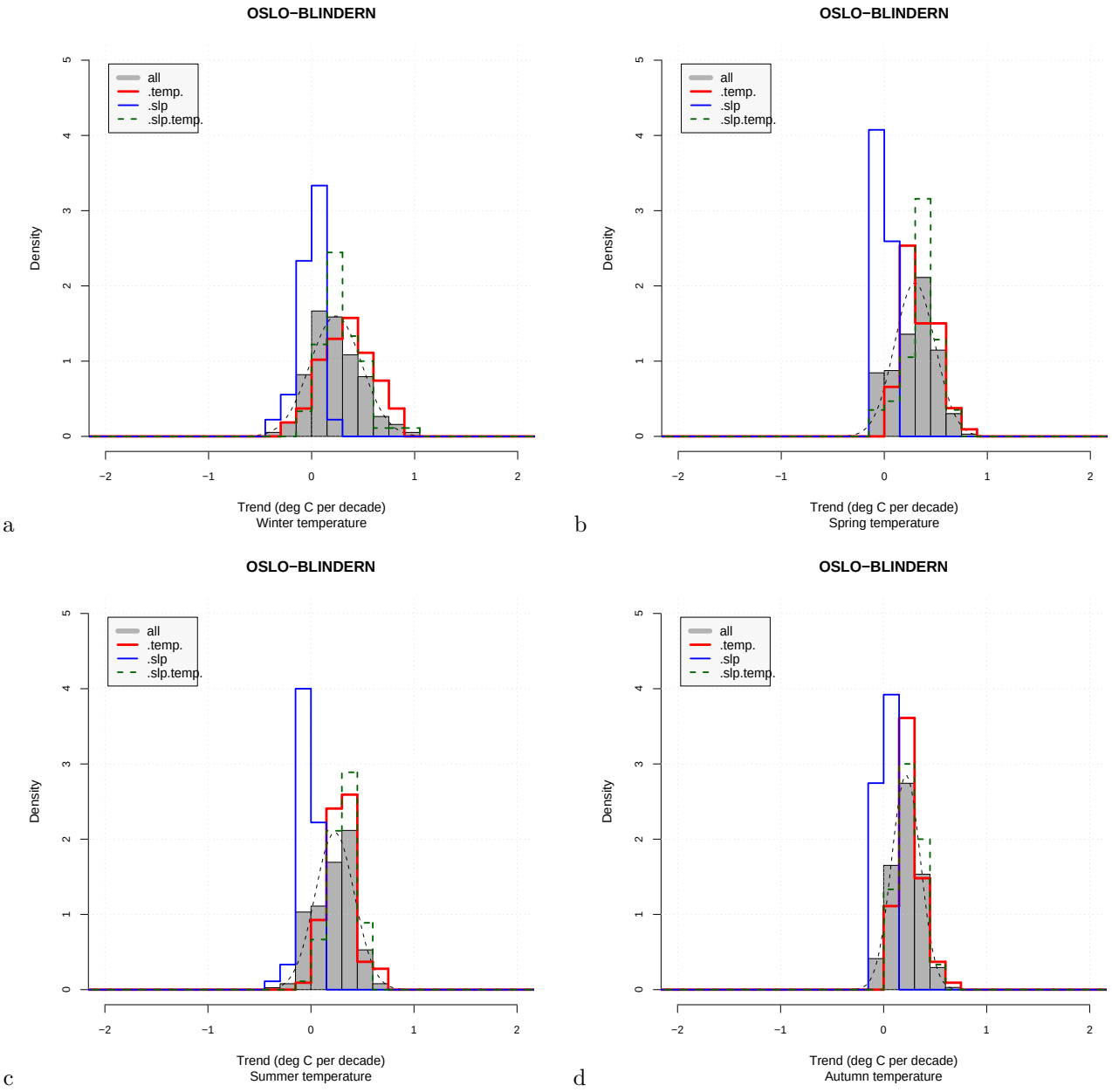

Figure 15. Histograms of downscaled monthly mean temperature trends for Oslo-Blindern presented in Figure 11 for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The thin black dashed line shows the best-fit Gaussian distribution of all the trend estimates and the red, blue, and green curves show histogram for subsets derived using temperature, SLP and a combination of temperature and SLP respectively.

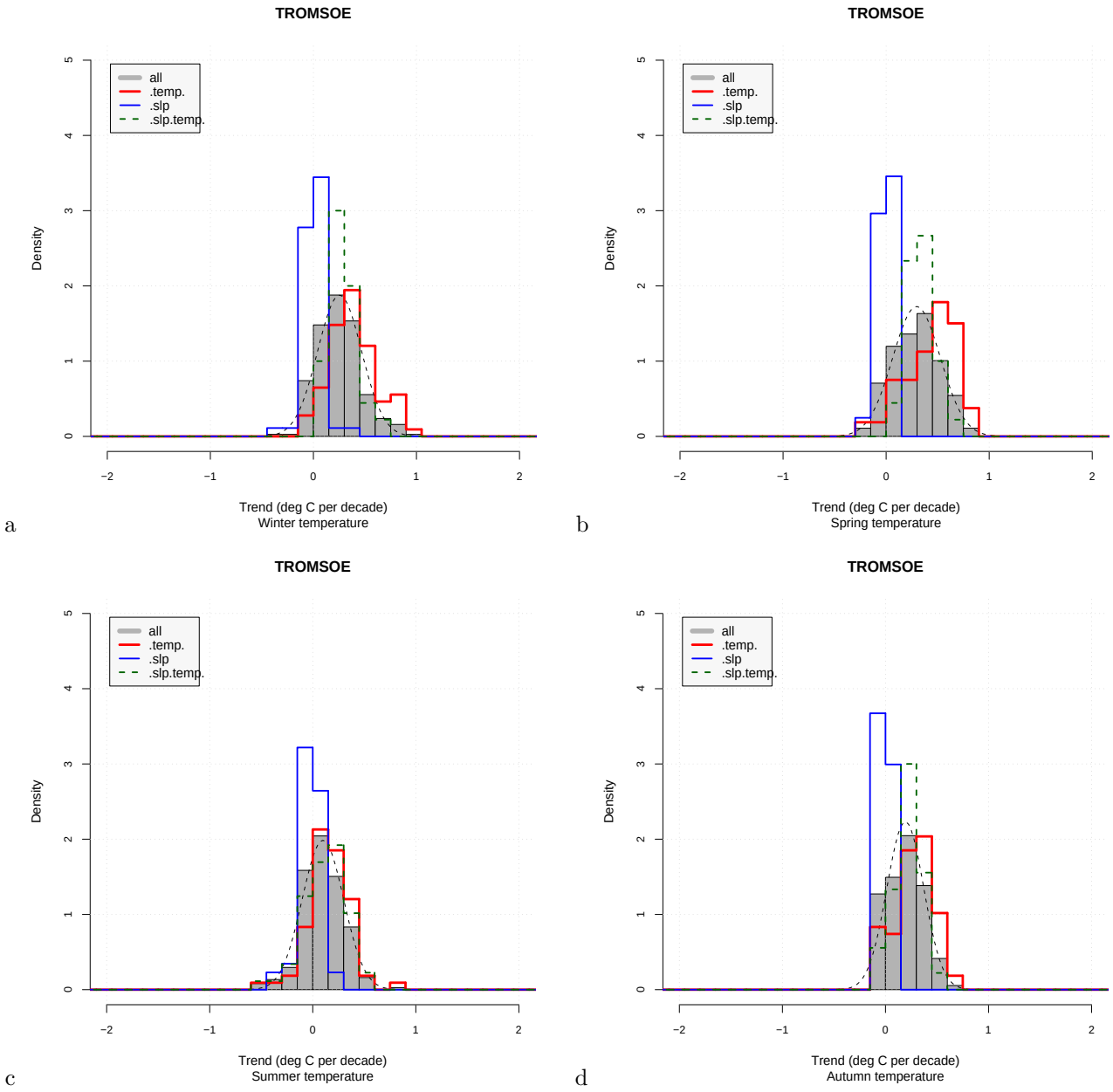

Figure 16. Histograms of downscaled monthly mean temperature trends for Tromsø presented for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The thin black dashed line shows the best-fit Gaussian distribution of all the trend estimates and the red, blue, and green curves show histogram for subsets derived using temperature, SLP and a combination of temperature and SLP respectively.

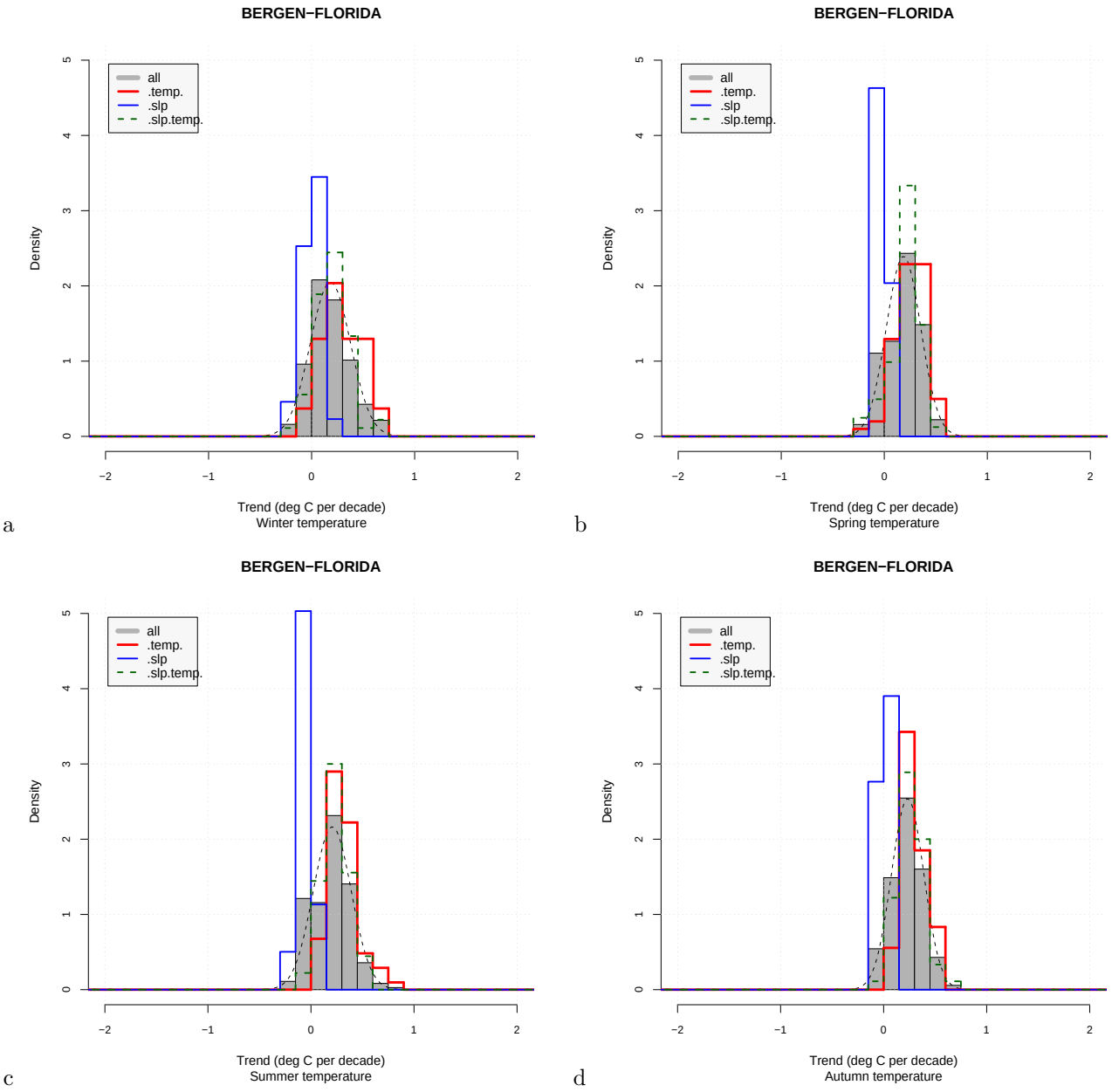

Figure 17. Histograms of downscaled monthly mean temperature trends for Bergen-Florida for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The thin black dashed line shows the best-fit Gaussian distribution of all the trend estimates and the red, blue, and green curves show histogram for subsets derived using temperature, SLP and a combination of temperature and SLP respectively.

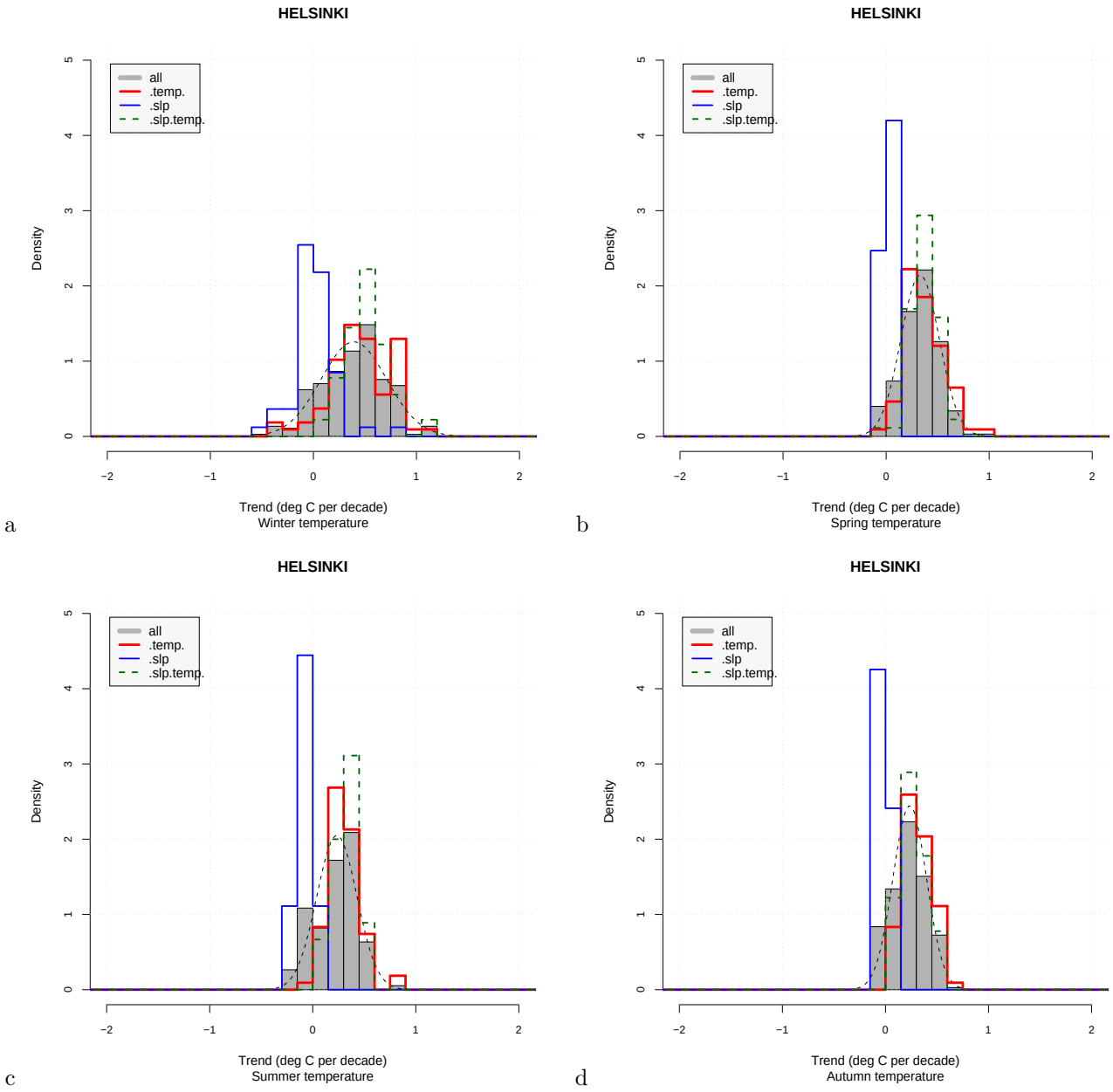

Figure 18. Histograms of downscaled monthly mean temperature trends for Helsinki for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The thin black dashed line shows the best-fit Gaussian distribution of all the trend estimates and the red, blue, and green curves show histogram for subsets derived using temperature, SLP and a combination of temperature and SLP respectively.

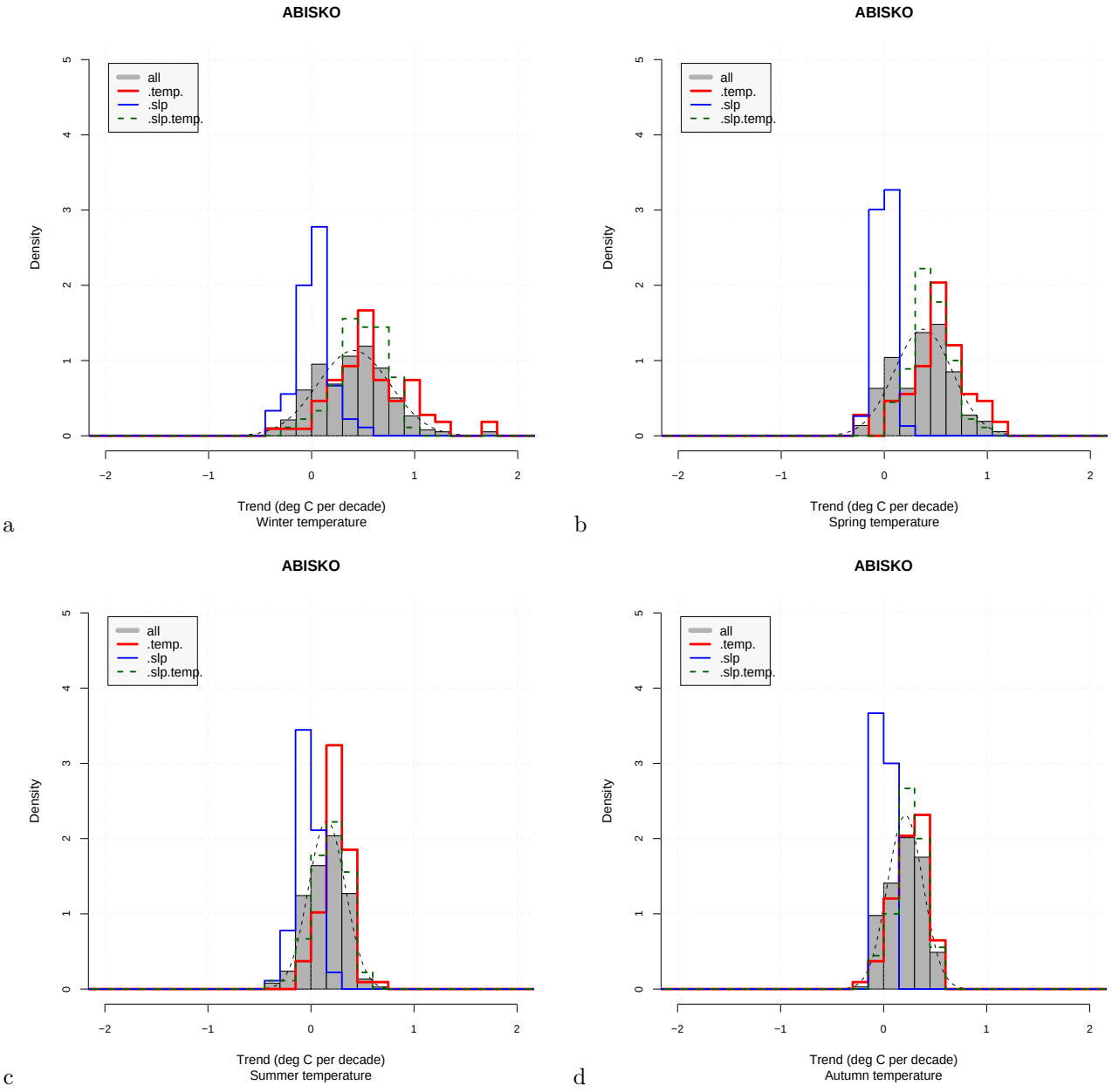

Figure 19. Histograms of downscaled monthly mean temperature trends for Abisko for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The thin black dashed line shows the best-fit Gaussian distribution of all the trend estimates and the red, blue, and green curves show histogram for subsets derived using temperature, SLP and a combination of temperature and SLP respectively.

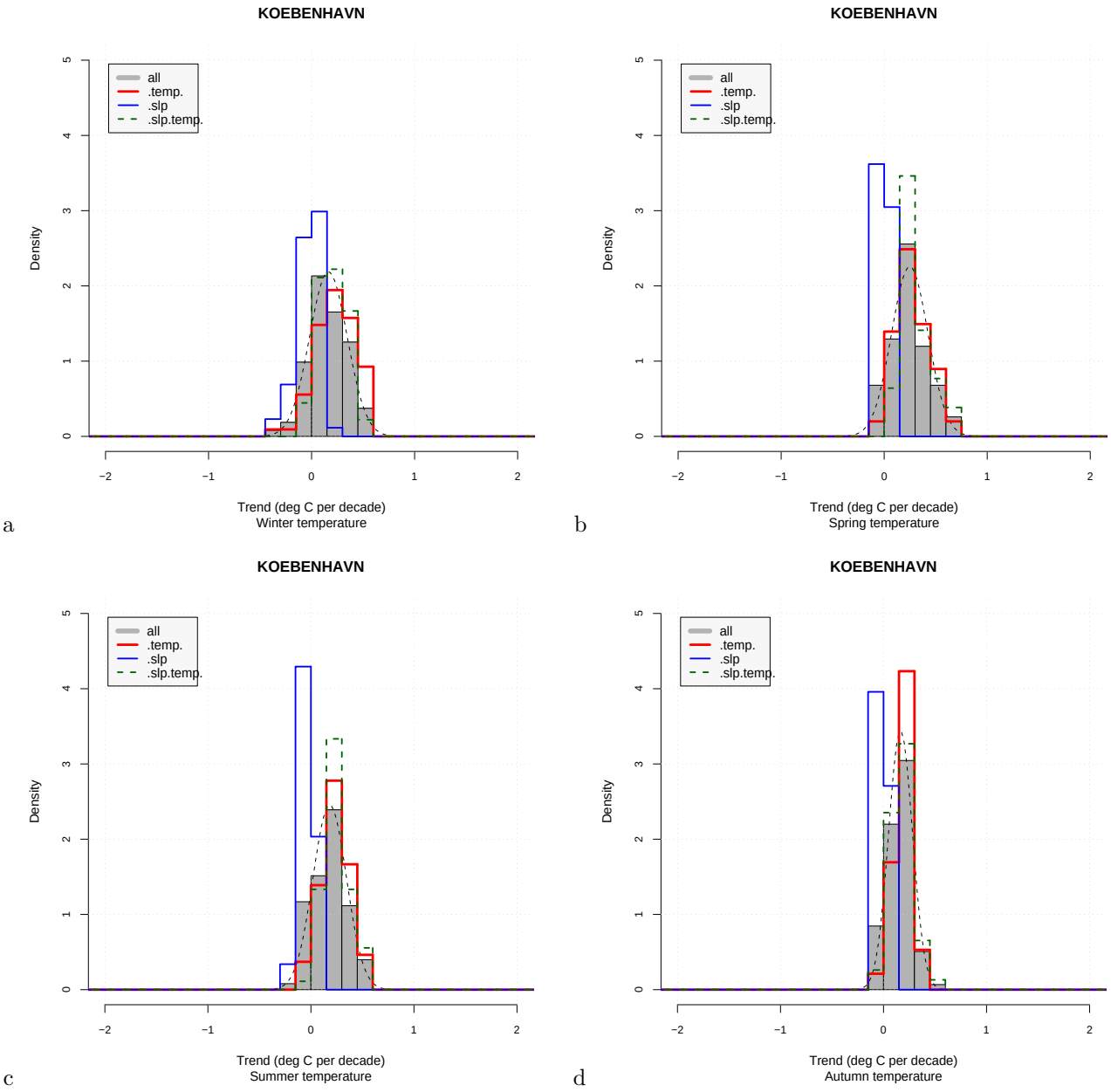

Figure 20. Histograms of downscaled monthly mean temperature trends for Copenhagen for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The thin black dashed line shows the best-fit Gaussian distribution of all the trend estimates and the red, blue, and green curves show histogram for subsets derived using temperature, SLP and a combination of temperature and SLP respectively.

3.5 Precipitation from the GCMs

3.5.1 The seasonality of precipitation

Large-scale fields of precipitation were used as predictor for the local rainfall (Figure 8(c,d)) and the quality of the trend estimates for these depends on whether the GCMs can reproduce realistically the large-scale prediction. One can assess the GCMs' ability to reproduce the annual variations in the precipitation in order to get an idea about the confidence that can be attached to the GCM output. Figure 21 shows the annual cycle of precipitation interpolated from the NCEP reanalysis and the various GCMs. Station-based observations (thick solid black line) are also shown to provide a reference to the actual precipitation. The solid grey curve shows the precipitation from the NCEP reanalysis, and is more appropriate for model evaluation because it relates to larger spatial scales the same way as the GCMs do. It is important to keep in mind that 'point' observations, such as station-based measurements, are not directly comparable to grid-box values from weather models. Large inter-model differences in the description of the seasonal cycle in precipitation is evident. In Bergen (a), the NCEP reanalysis gives a similar annual cycle as the station measurements at Bergen-Florida, however, some GCMs tend to get the annual peak too late in the autumn. Most GCMs tend to overestimate the mean precipitation. At Falun (b), the NCEP-reanalysis gives substantially stronger annual peaks than the in-situ observations.A number of GCMs get the annual timing of the precipitation peak right, but there are also GCMs that predicts the peak too late in the autumn. Neither the NCEP-reanalysis nor any of the GCMs reproduce the magnitude of the precipitation recorded at Glomfjord (c). Most GCMs give a reasonable description of the magnitude of rain fall annual cycle in Helsinki (d), albeit a few months later than the actual peak. One exception is the CCCma model that produces a seasonal cycle that is anti-correlated with the observations. Similar results for Copenhagen, Oslo, Stockholm and Tromsø are shown in Figure 22. The ECHAM4/OPYC3 and the NCAR-PCM models produce completely unrealistic results for Copenhagen (a), and there are no GCM that reproduces the annual cycle for this location. The NCEP-reanalysis has a bias towards lower values, but otherwise a realistic seasonal cycle. Once more, the CCCma model gave unrealistic results. Most of the GCMs produced a phase-lag in the annual cycle for Oslo (b) but the magnitude of the cycle was in general realistic, albeit with an offset in the mean level. The GCM results interpolated to Stockholm (c) were characterised by phase lag and unrealistic magnitudes. Even the NCEP re-analysis indicated a phase lag and a low bias. Part of these discrepancies may be related to inappropriate interpolation, e.g. where grid boxes over say the ocean have too much influence for the interpolated location. As in Copenhagen, the CCCma produced unrealistic results. In Tromsø, most GCMs got the phase of the annual cycle right, albeit with too weak magnitude. In this case, the GCMs seemed to produce a more realistic seasonal cycle than the NCEP re-analysis. Again, the CCCma distinguishes itself from the other GCMs by producing a seasonal cycle that is anti-correlated with the observations.

The large differences between the annual precipitation cycle interpolated from the various GCMs, and the large discrepancies regarding the observed and re-analysed data, may have serious implications. If the GCMs are not able to describe how the precipitation responds to annual variations in the boundary forcing, then how can these models be expected to account for how the preciptation will change with increased atmospheric consentrations of greenhouse gases? In order to answer this question, one must examine the seasonal cycle of the observations to see if it is robust and well-defined or whether the annual variations in the boundary conditions have a weaker influence on the precipitation in the in-land locations. Figure 23 shows the spread about the annual cycle in Oslo and Copenhagen, and it is clear from this plot that (chaotic) internal variations in these locations are more prominent than the influence from the annual variations in the short-wave radiation. For Bergen, on the other hand, the good correspondence between the modelled and observed seasonal variations suggest a strong external influence on the precipitation on the west coast of Norway where the wind direction and orographic influences play an important role. Hence, the lack of consistency in describing the seasonal precipitation cycle amongst the GCMs doesn't mean that these models necessarily are flawed, but it highlights the difficulty predicting future changes in the precipitation. This interpretation is in accord with Ruosteenoja et al. (2003), who found substantial differences between modeled and observed annual cycle in precipitation over sub-continental scale regions (10^6–10^8km^2). They also suggested that part of the model-observation discrepancies may be due to stronger inter-annual variations in precipitation and stronger influence of small-scale geographical features on local precipitation.

Hanssen-Bauer et al. (2003) and Goodess et al. (2003) found that a more realistic annual cycle can be obtained through empirical downscaling. An evaluation based on the `clim.pact` support their finding. Figure 24 shows the empirically downscaled monthly precipitation for the whole year for Oslo-Blindern and Copenhagen. The top panels show the time series of full values whereas the lower panels show the seasonal precipitation cycle. The empirical downscaling ensures a realistic annual cycle since it produces anomalies that are superimposed onto the observed annual cycle. The reasonably high skill-scores ($R^2 > 50\%$) suggests that the anomalies are well described. This example shows the problem taking precipitation directly from GCMs in order to describe local precipitation.

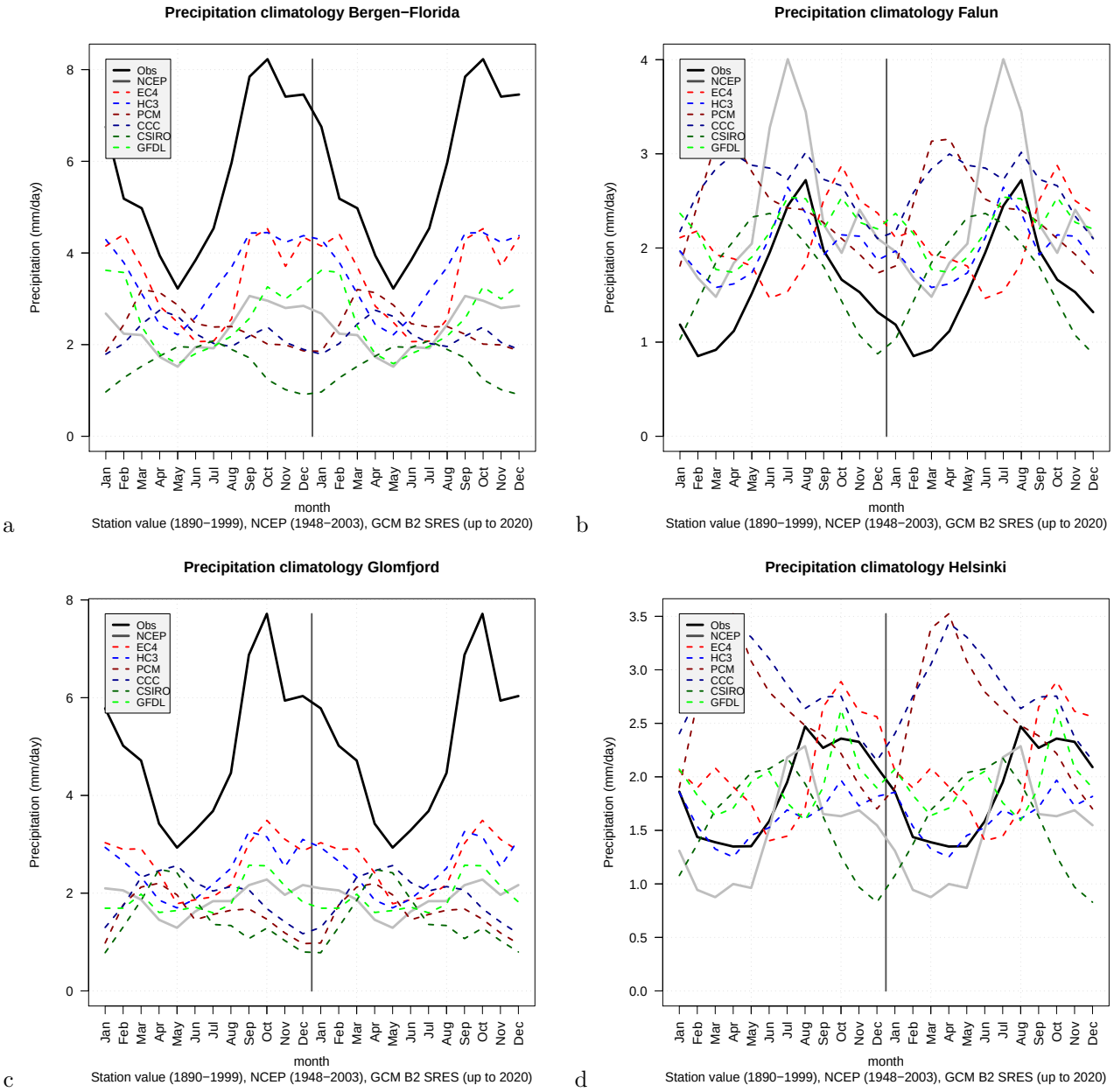

Figure 21. A comparison between the annual cycle in the precipitation for Bergen, Falun, Glomfjord and Helsinki. The station-based measurements (black) are shown together with the NCEP-reanalysis and bi-linearly interpolated estimates for various GCMs. The time interval varies for the different GCMs depending on the starting model date of the integration ()

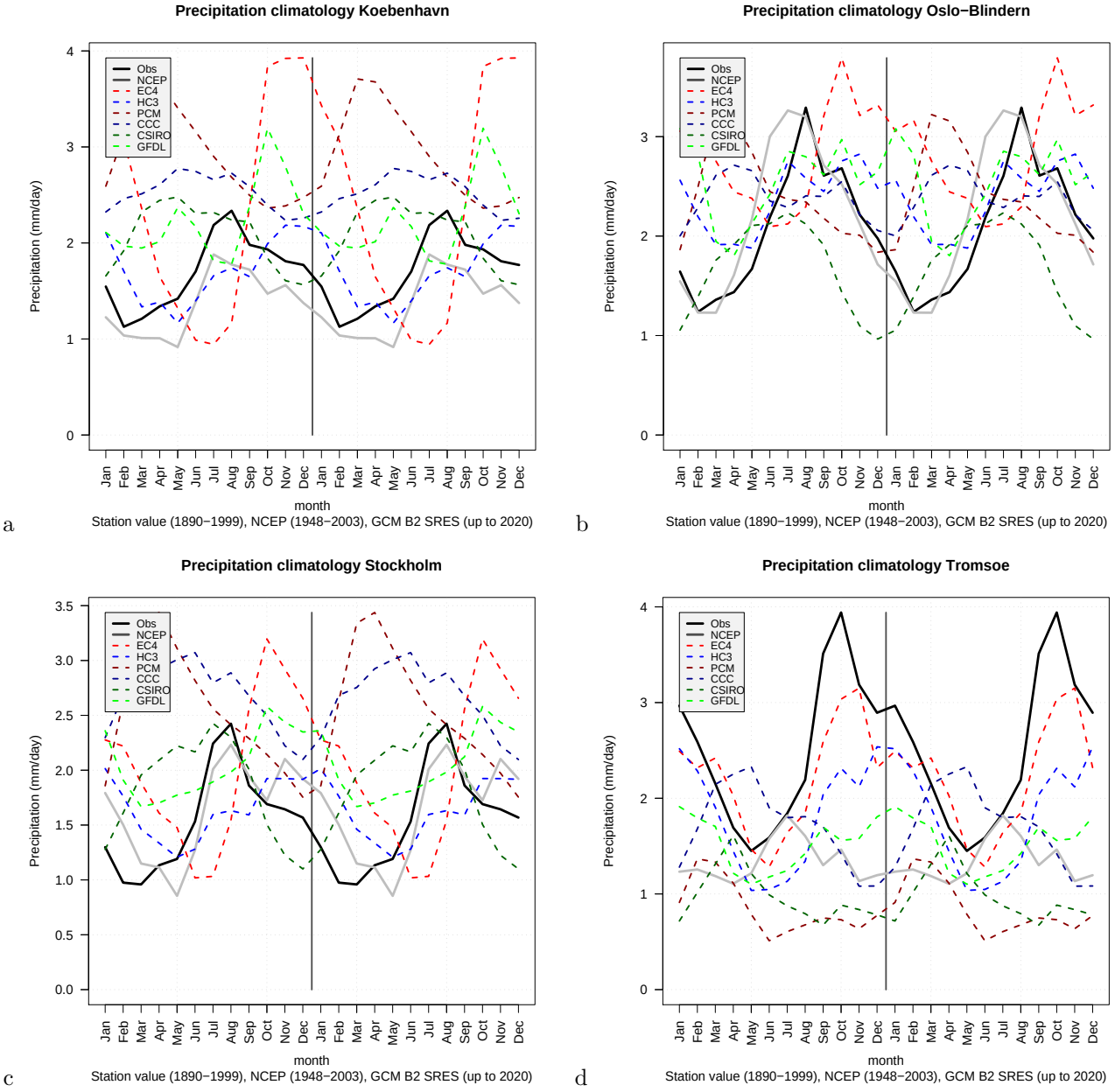

Figure 22. A comparison between the annual cycle in the precipitation for Copenhagen, Oslo-Blindern, Stockholm and Tromsø. The station-based measurements (black) are shown together with the NCEP-reanalysis and bi-linearly interpolated estimates for various GCMs.

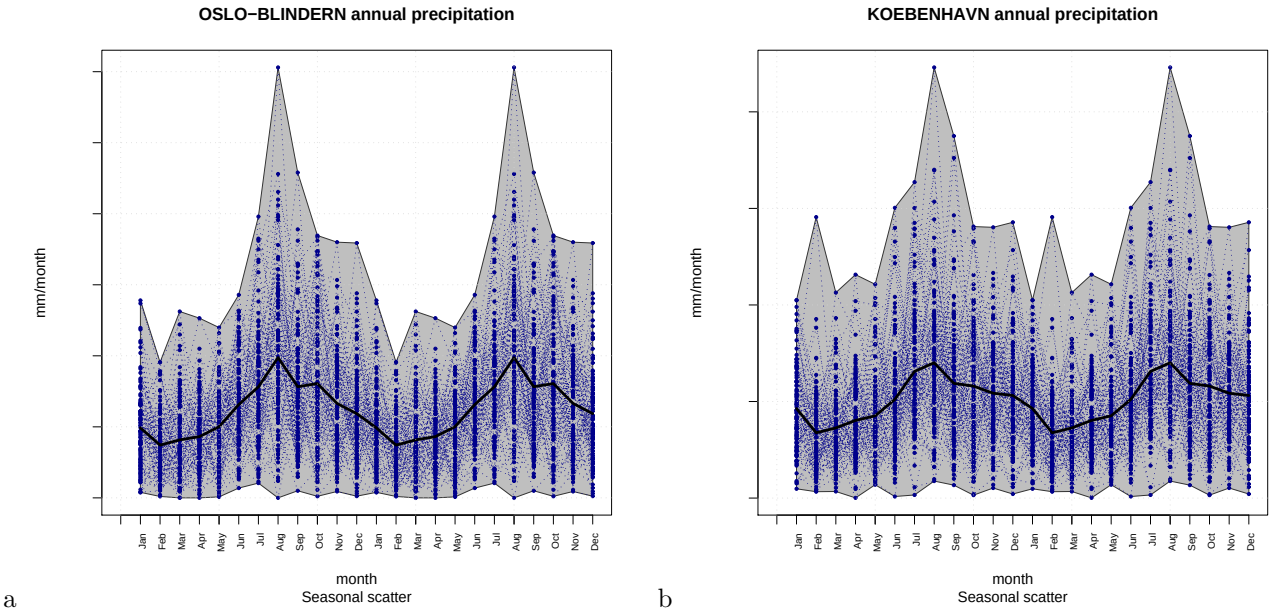

Figure 23. The substantial scatter about the seasonal variations in monthly precipitation at Oslo and Copenhagen suggests that the annual cycle is not well-defined.

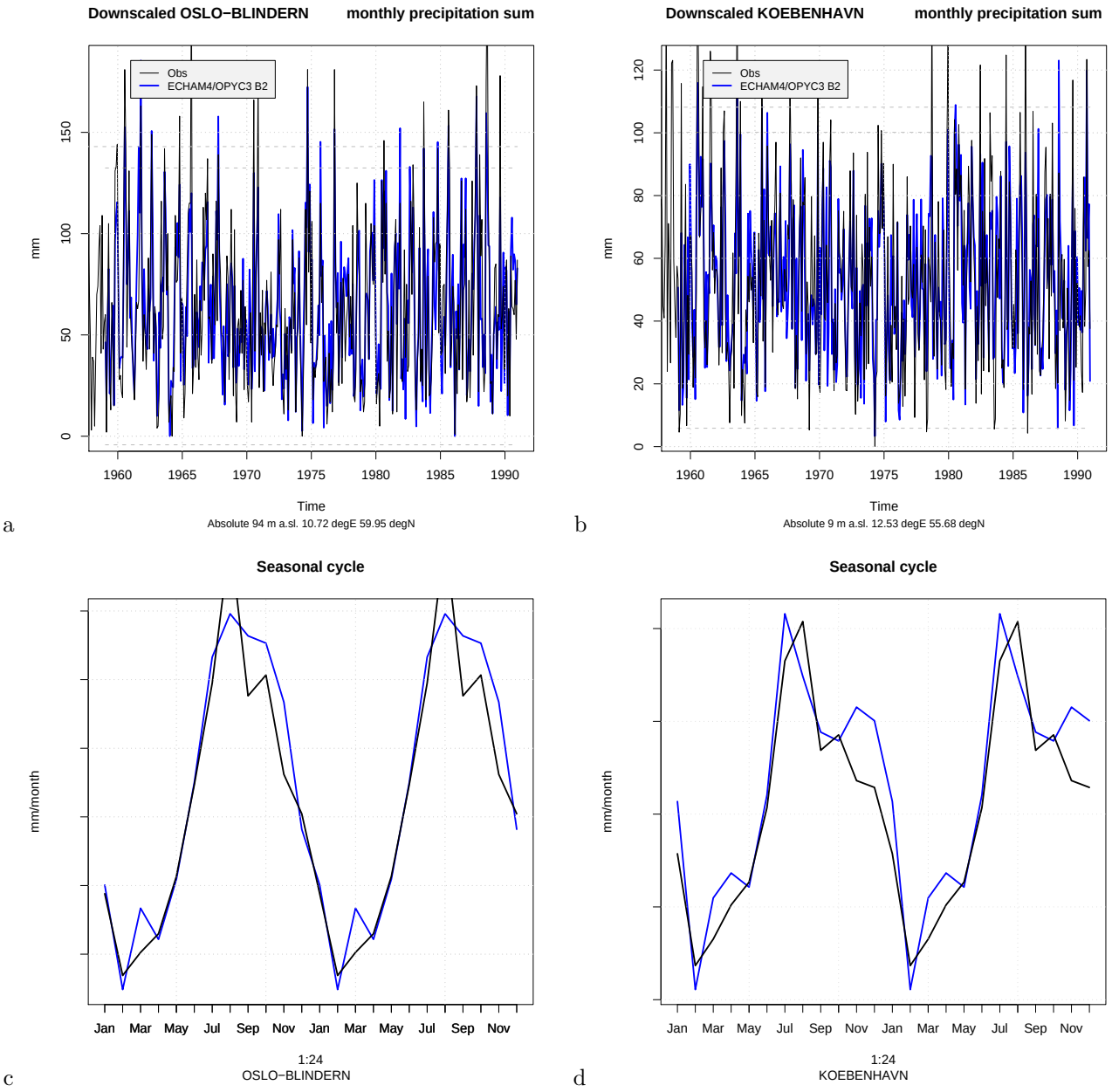

Figure 24. Evaluation of empirically downscaled precipitation for Oslo and Copenhagen based on NCEP large-scale precipitation fields. R^2-scores for Oslo precipitation: 78, 68, 53, 70, 53, 65, 64, 56, 60, 88, 80, 77% for the 12 Calendar months respectively, and for Copenhagen 61, 73, 69, 71, 58, 80, 52, 56, 69, 81, 64, 77%. Top: absolute values; bottom: seasonal cycle.

3.5.2 Large-scale precipitation anomalies

In order to examine the large-scale precipitation in more detail, the common EOF products based on the NCEP re-analysis and GCM precipitation were plotted in a similar fashion as the mixed-common EOFs in Figures 2 – 7. Figure 25 shows these diagnostics for January monthly mean precipitation in the ECHAM4/OPYC3 SRES B2 results, indicating maximum weights west off Scotland that also covers an extensive part of the North Sea and south of Iceland. The model magnitudes are similar to those in the NCEP re-analysis. Somewhat surprisingly, very similar spatial patterns and magnitude are seen in the HadCM3 and the NCAR-PCM results (Figure 27). The CCCma model (Figure 28) produces too weak variability in the precipitation, whereas the CSIRO results (Figure 29a) are consistent with the other GCMs'.

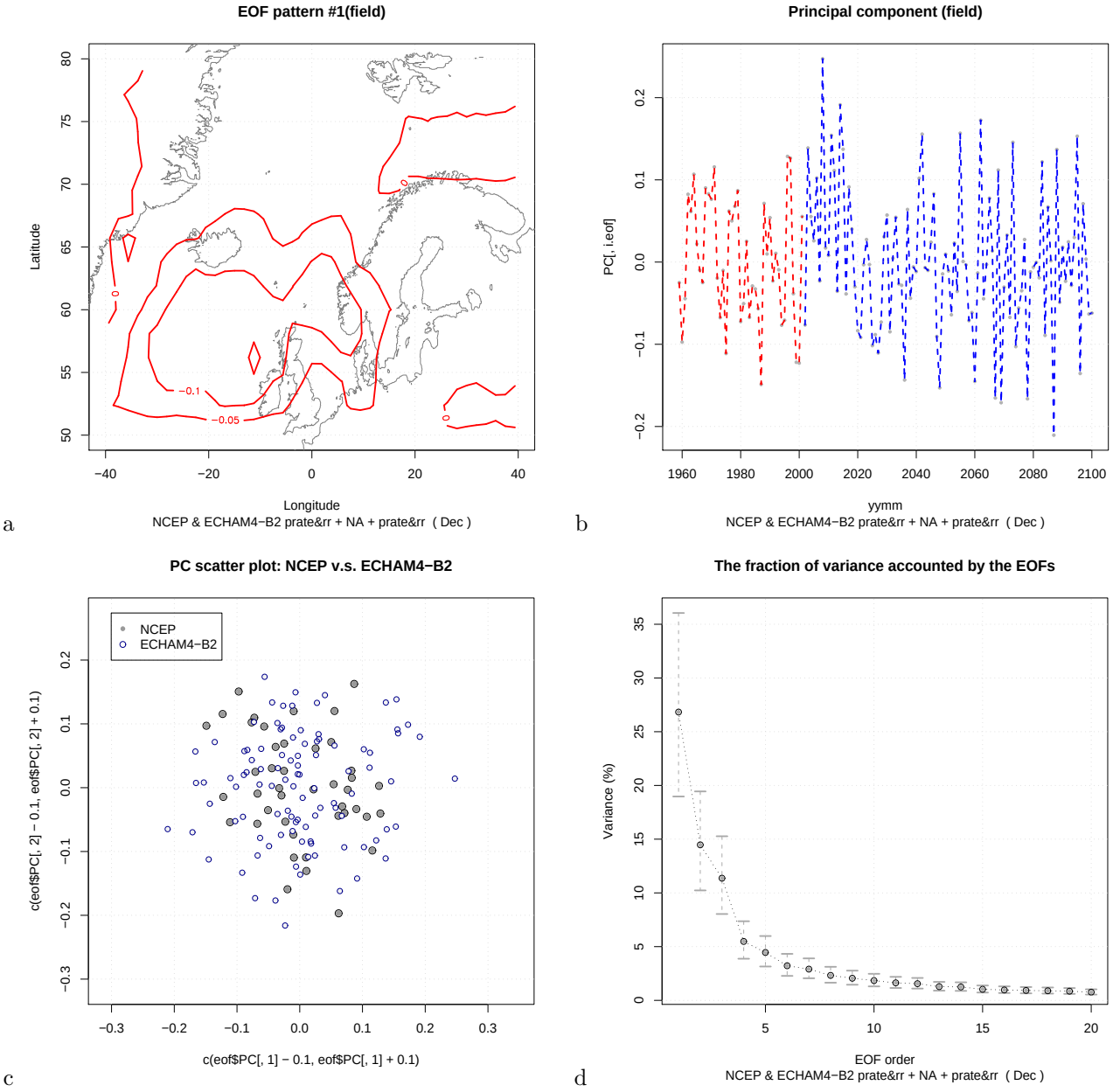

Figure 25. Common EOF products for large-scale precipitation from the ECHAM4/OPYC3 B2 integration for January. (a) leading spatial EOF patterns, (b) leading PC, (c) comparison between the scatter of the two leading PCs from NCEP reanalysis and the GCM, (d) the proportional variance described by the leading modes.

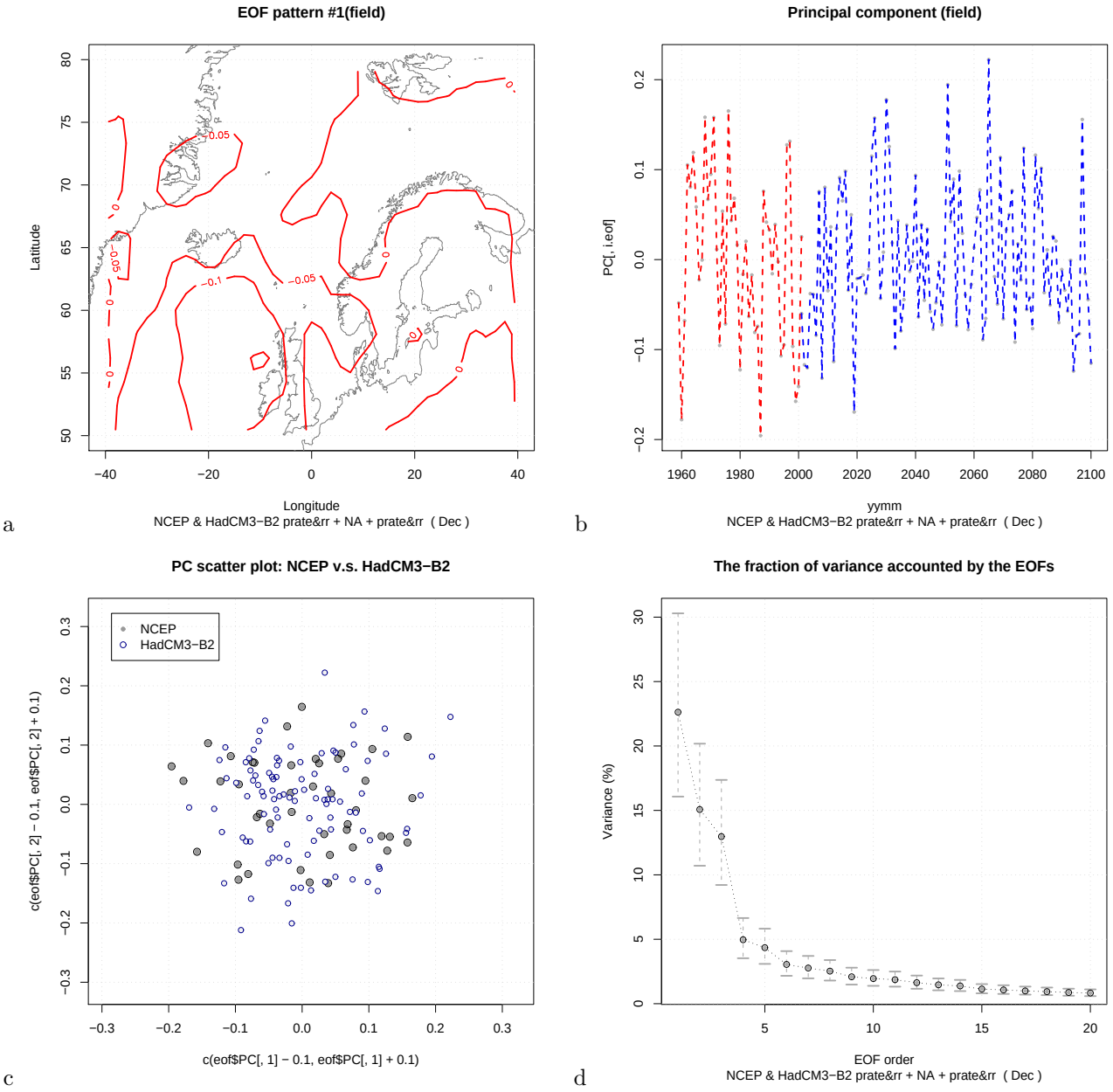

Figure 26. Common EOF products for large-scale precipitation from the HadCM3 B2 integration for January. (a) leading spatial EOF patterns, (b) leading PC, (c) comparison between the scatter of the two leading PCs from NCEP reanalysis and the GCM, (d) the proportional variance described by the leading modes.

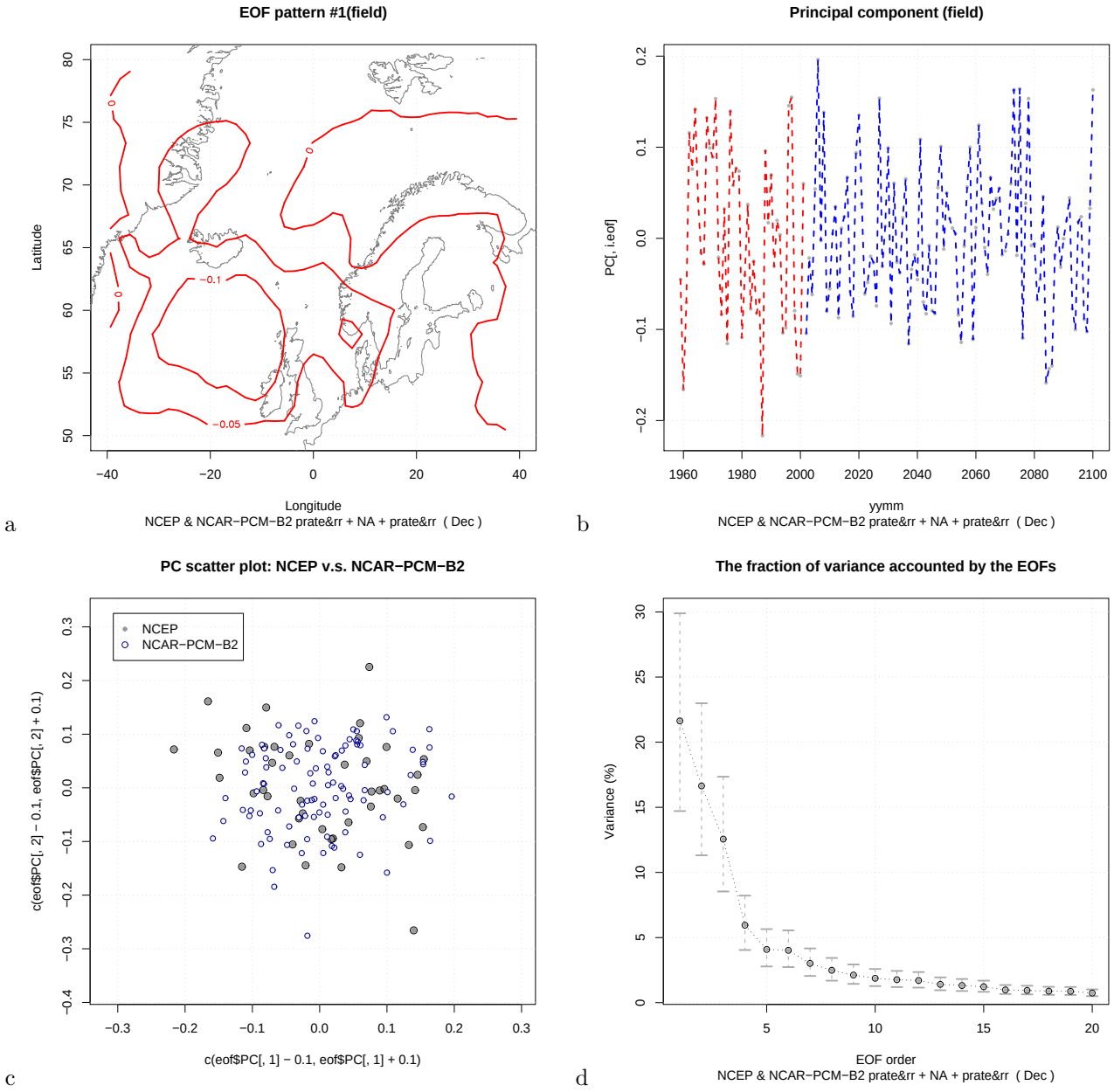

Figure 27. Common EOF products for large-scale precipitation from the NCAR-PCM B2 integration for January. (a) leading spatial EOF patterns, (b) leading PC, (c) comparison between the scatter of the two leading PCs from NCEP reanalysis and the GCM, (d) the proportional variance described by the leading modes.

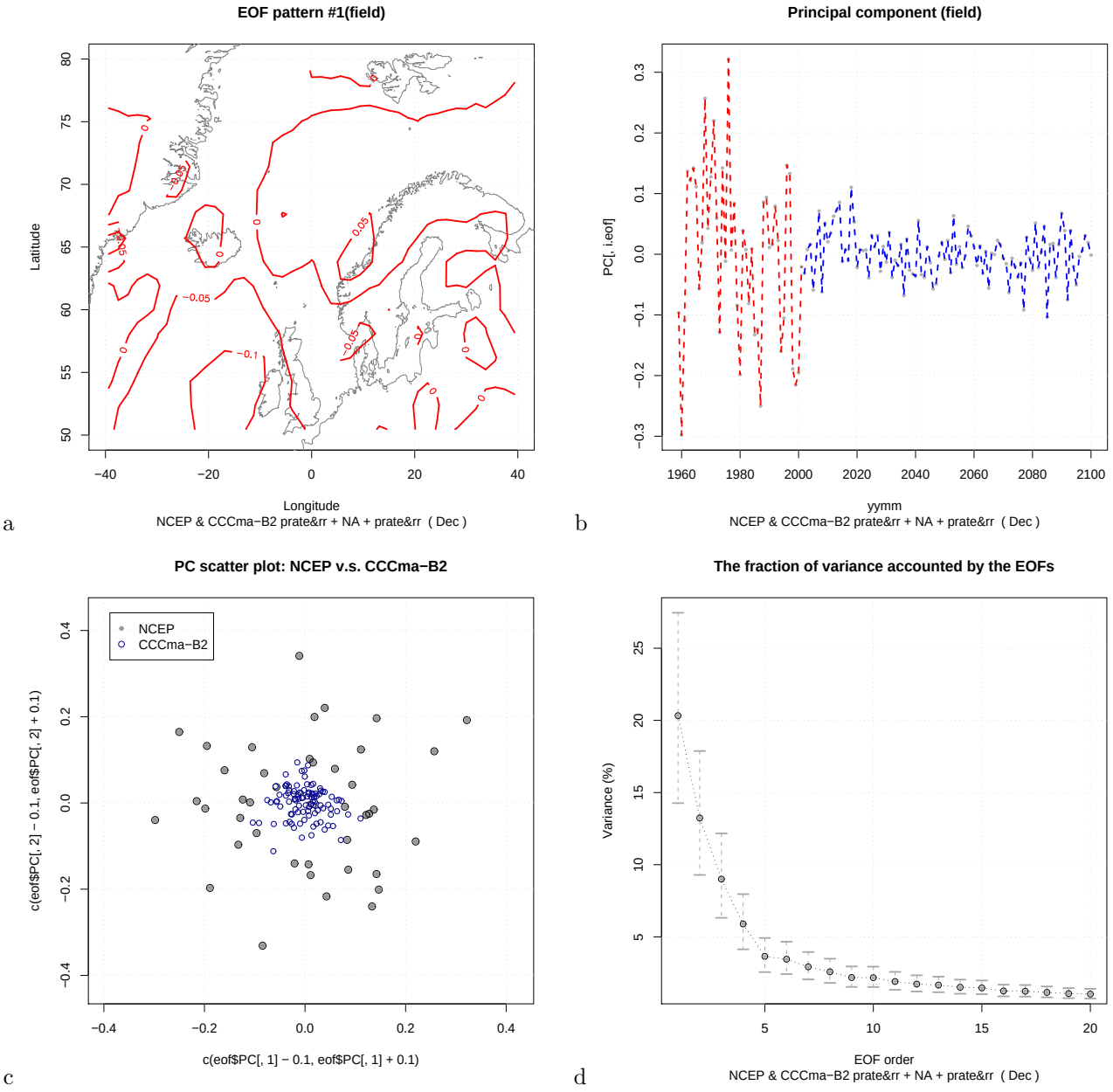

Figure 28. Common EOF products for large-scale precipitation from the CCCma B2 integration for January. (a) leading spatial EOF patterns, (b) leading PC, (c) comparison between the scatter of the two leading PCs from NCEP reanalysis and the GCM, (d) the proportional variance described by the leading modes.

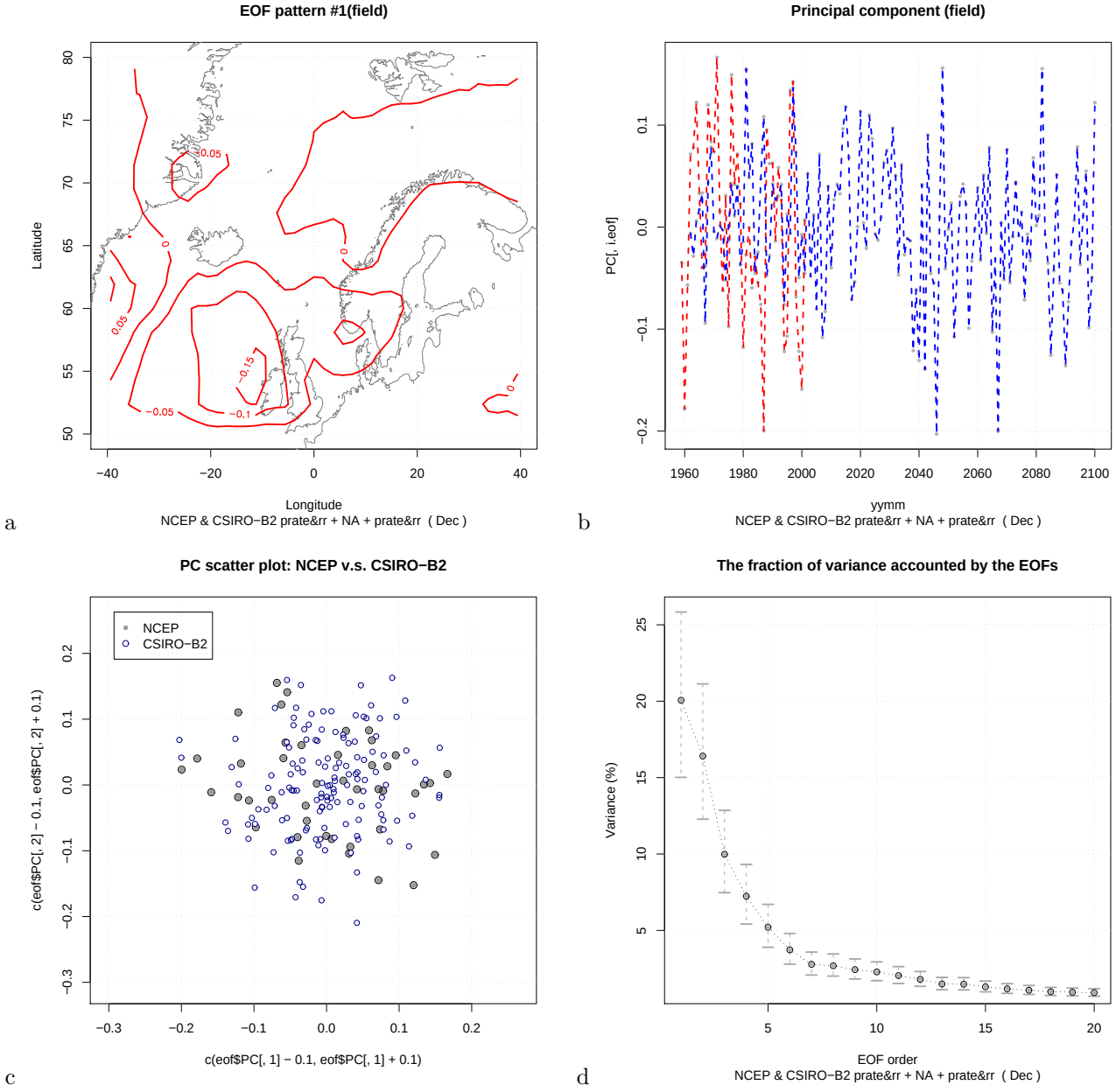

Figure 29. Common EOF products for large-scale precipitation from the CSIRO B2 integration for January. (a) leading spatial EOF patterns, (b) leading PC, (c) comparison between the scatter of the two leading PCs from NCEP reanalysis and the GCM, (d) the proportional variance described by the leading modes.

3.6 Spectral properties of large-scale precipitation anomalies

A wavelet analysis (Addison, 2004; Torrence & Compo, 1998) was applied to the leading PCs in order to compare and investigate the spectral characteristics of the Dec precipitation anomalies in the GCMs with reference to the NCEP reanalysis (Figure 30–31. The ECHAM4/OPYC3 model described more pronounced low-frequency variability in the precipitation, especially toward the end of the integration (e.g. for the future). HadCM3 and NCAR-PCM produced similar spatial characteristics as seen in the NCEP December precipitation. For June, the NCEP precipitation exhibits more pronounced low-frequency fluctuations, and none of the GCMs give similar spectral characteristics. The HadCM3 results are closest to the NCEP in terms of June spectral properties, as this GCM produced strongest low-frequencies of the GCMs.

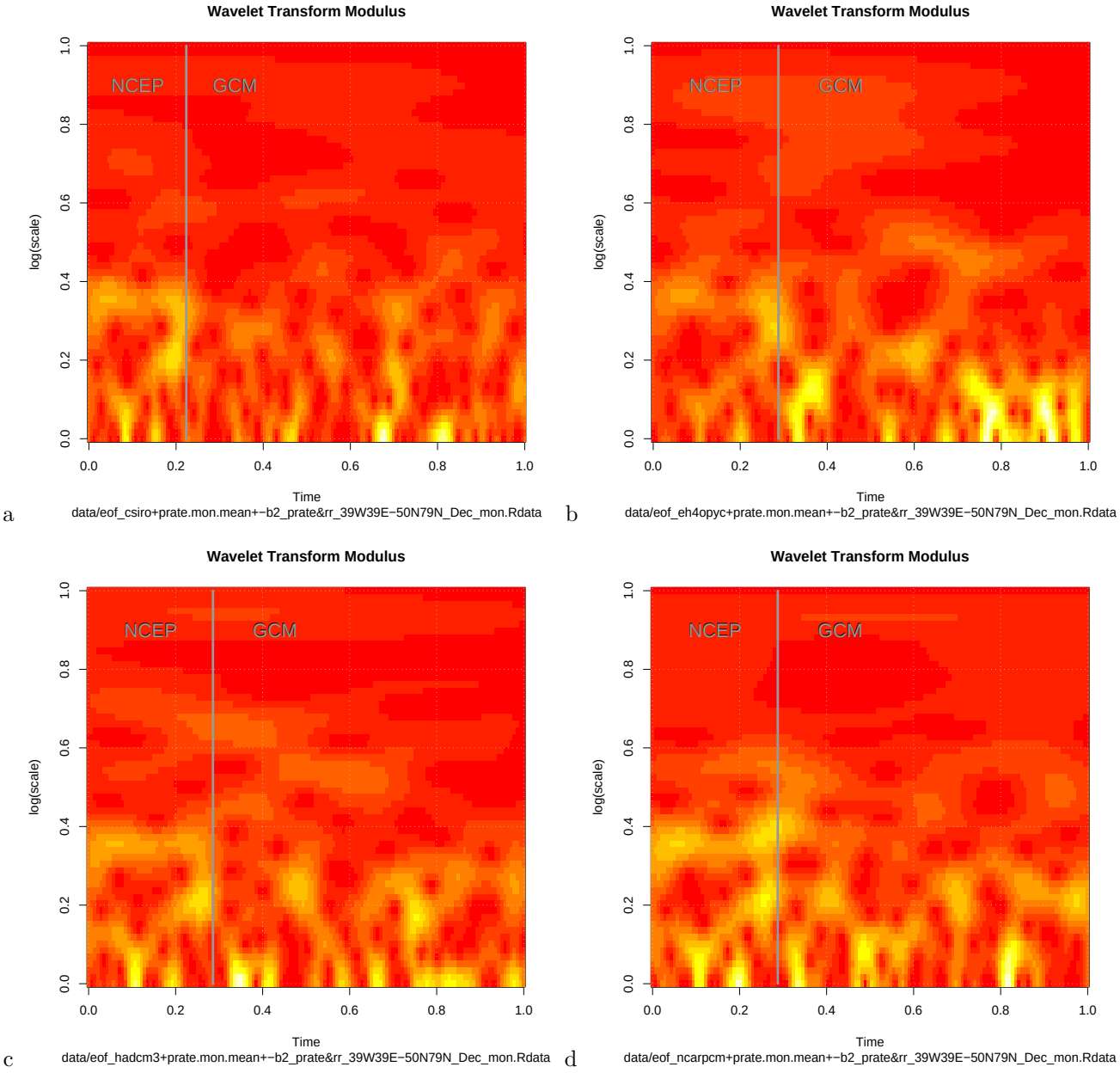

Figure 30. Wavelet analysis of the leading common PC (Using `Rwave: cwt()`) for the December month: a) CSIRO, b) ECHAM4/OPYC3, c) HadCM3, and d) NCAR-PCM. The wavelet analysis for CCCma is not shown.

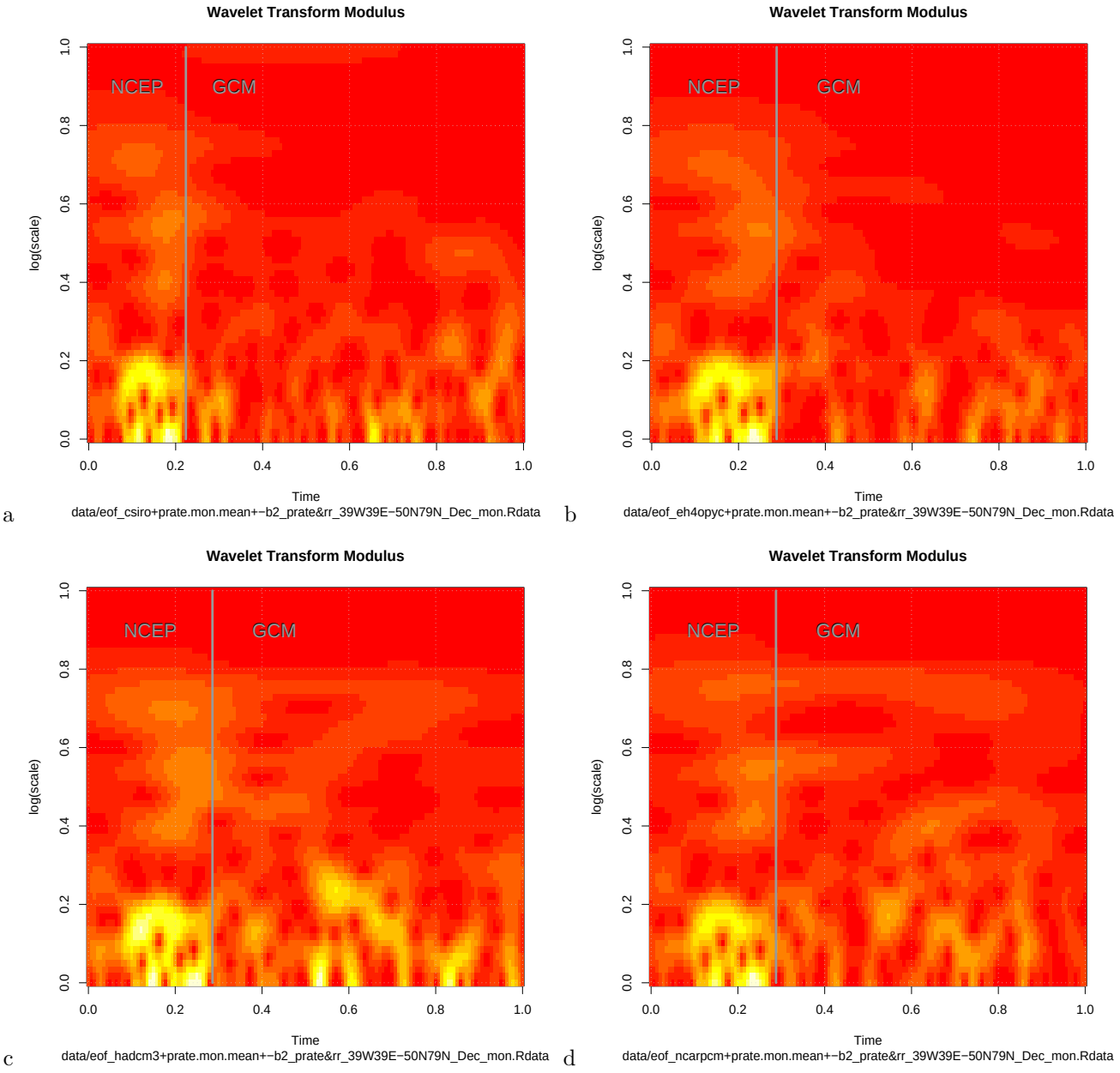

Figure 31. Wavelet analysis of the leading common PC (Using `Rwave: cwt()`) for the June: a) CSIRO, b) ECHAM4/OPYC3, c) HadCM3, and d) NCAR-PCM. The wavelet analysis for CCCma is not shown.

3.7 Downscaled precipitation trends

Figure 32 shows a similar analysis as shown in Figure 11, but for precipitation instead of temperature. There is little evidence of a common signal: most of the boxes straddle the zero-line. The largest scatter is found in summer. There may be a weak hint of a bias towards lower precipitation in spring and summer, although not significant, and four out of six GCMs give trends toward wetter autumn conditions.

Figure 33 presents a summary of the winter precipitation at Tromsø, Stykkisholmur, Stockholm and Abisko. For Tromsø (a), there are substantial inter-model differences: the GFDL and the HadCM3 models produce much smaller range of values and these are clearly well below the zero-line. Four out of the six GCMs indicate a bias toward wetter future autumn conditions. For Stykkisholmur (b), most

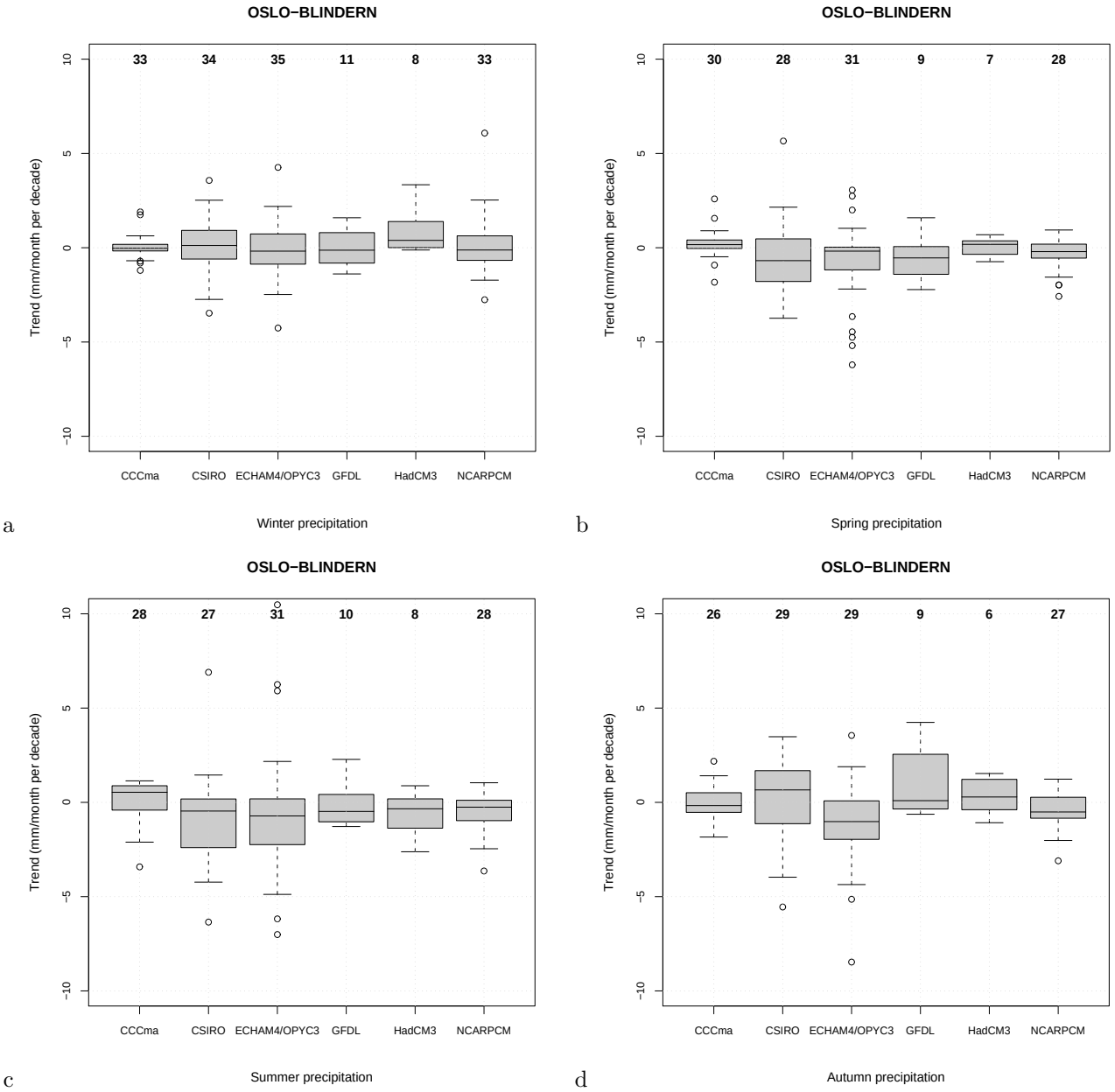

Figure 32. Downscaled monthly mean precipitation trends for Oslo-Blindern for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The number of data points in each box is denoted near the top axis.

of the GCMs indicate a bias towards positive precipitation trends, and it is only the HadCM3 that points to dryer conditions. The HadCM3 model produces unrealistically cold conditions for the present-day climate in this region (Benestad et al., 2002), and the reliability of these estimates are therefore questionable. There is little discernable signal in the trends for winter precipitation in Stockholm (c) and Abisko (d).

Figure 34 is similar to Figure 15 but for precipitation in Oslo rather than temperature. The histograms indicate that the distributions peak near zero-trend, although there are hints of secondary peaks at negative values in summer and positive values in autumn. When the precipitation rate is used as a predictor (red), then the results tend to point to wetter trends than when the SLP is used. Figures 35–39 show similar analysis for Tromsø, Bergen-Florida, Helsinki, Abisko, and Copenhagen. The

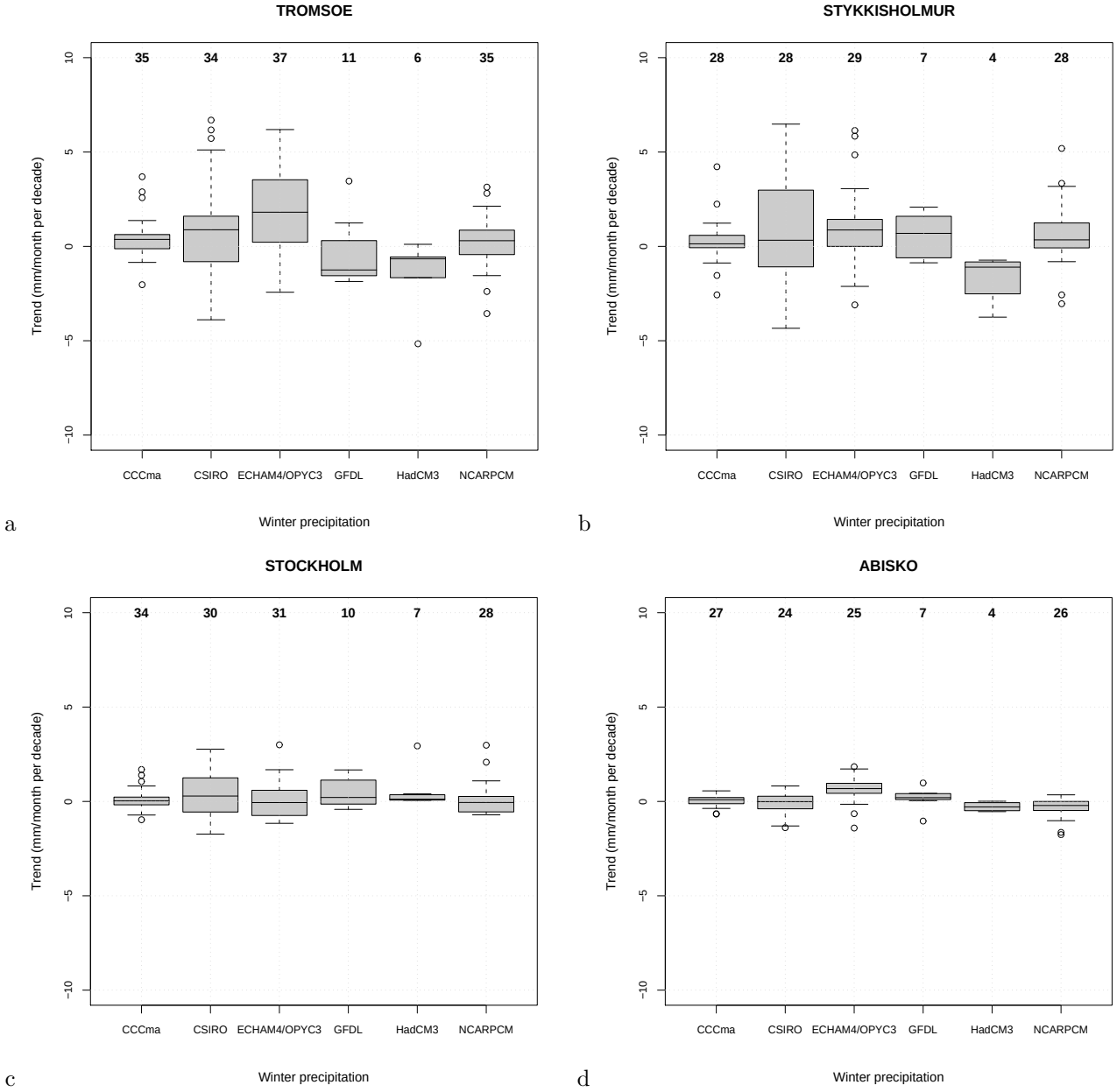

Figure 33. Downscaled monthly mean precipitation trends for December–February in Tromsø (a), Stykkisholmur (b), Stockholm (c) and Abisko (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included. The number of data points in each box is denoted near the top axis.

SLP-predictor gave wetter winter-estimates for Tromsø than predictors consisting of large-scale precipitation (Figure 36a). The spread reflects the range of values, with the widest distribution in Bergen and narrow distributions in Abisko and Copenhagen. These differences in spread is expected and due to the differences in the total precipitation amounts. In general,there was no clear consensus between the GCMs as to whether one can expect a wetter or drier future climate if the derivation of local precipitation was based on both SLP and large-scale precipitation rate. If, on the other hand, the downscaling only included the latter predictor, then the results point to a tendency of wetter climates in the future for Oslo (winter and spring), Tromsø (spring), Bergen, Helsinki, Abisko and Copenhagen (all seasons). The reason for the systematic difference between trends derived from SLP only and those from the other predictors may be that a gradual global warming doesn't influence the SLP field in a way that can bee seen in terms of local temperatures and precipitation.

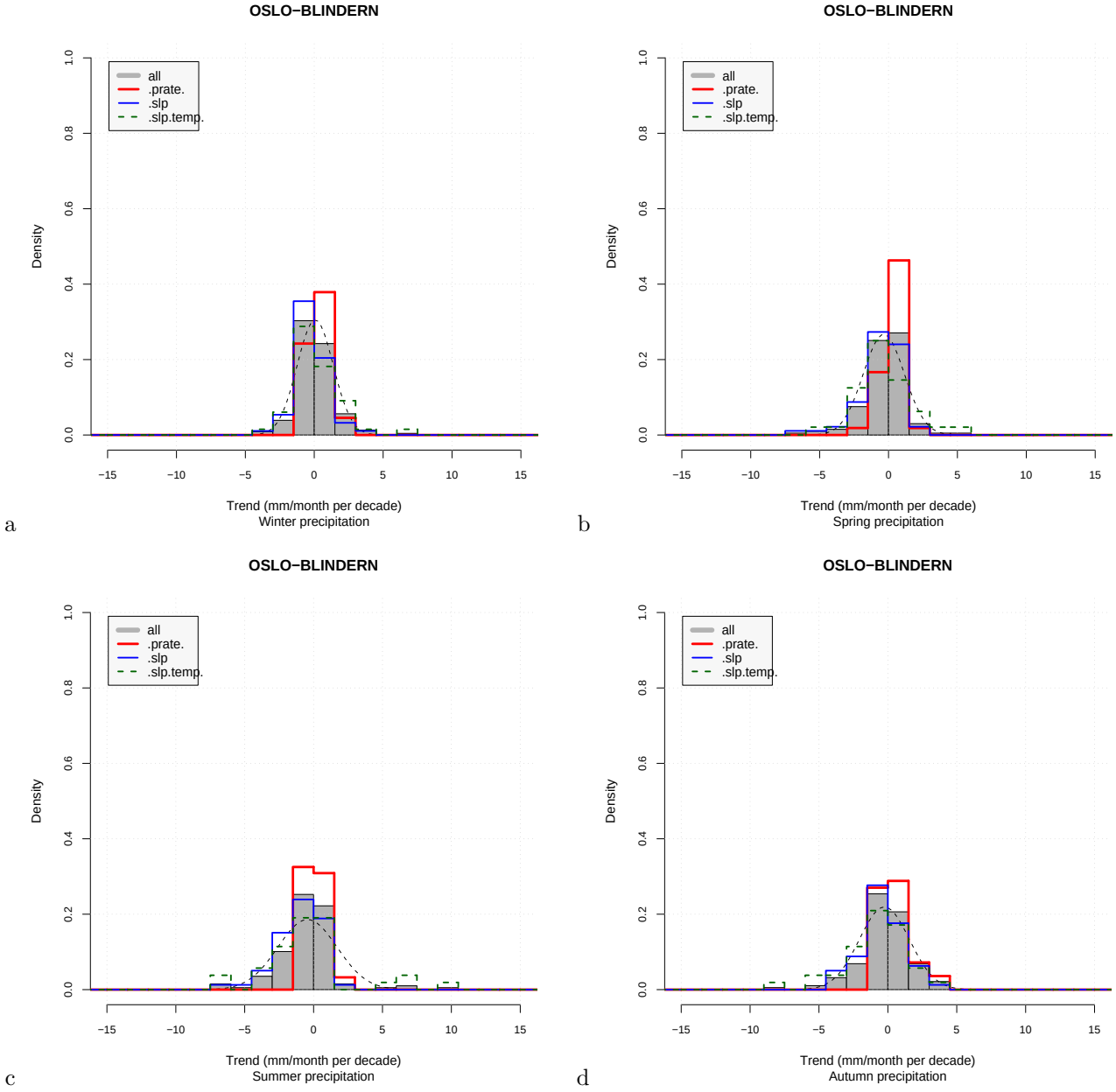

Figure 34. Histograms of downscaled monthly mean precipitation trends for Oslo-Blindern presented in Figure 11 for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included.

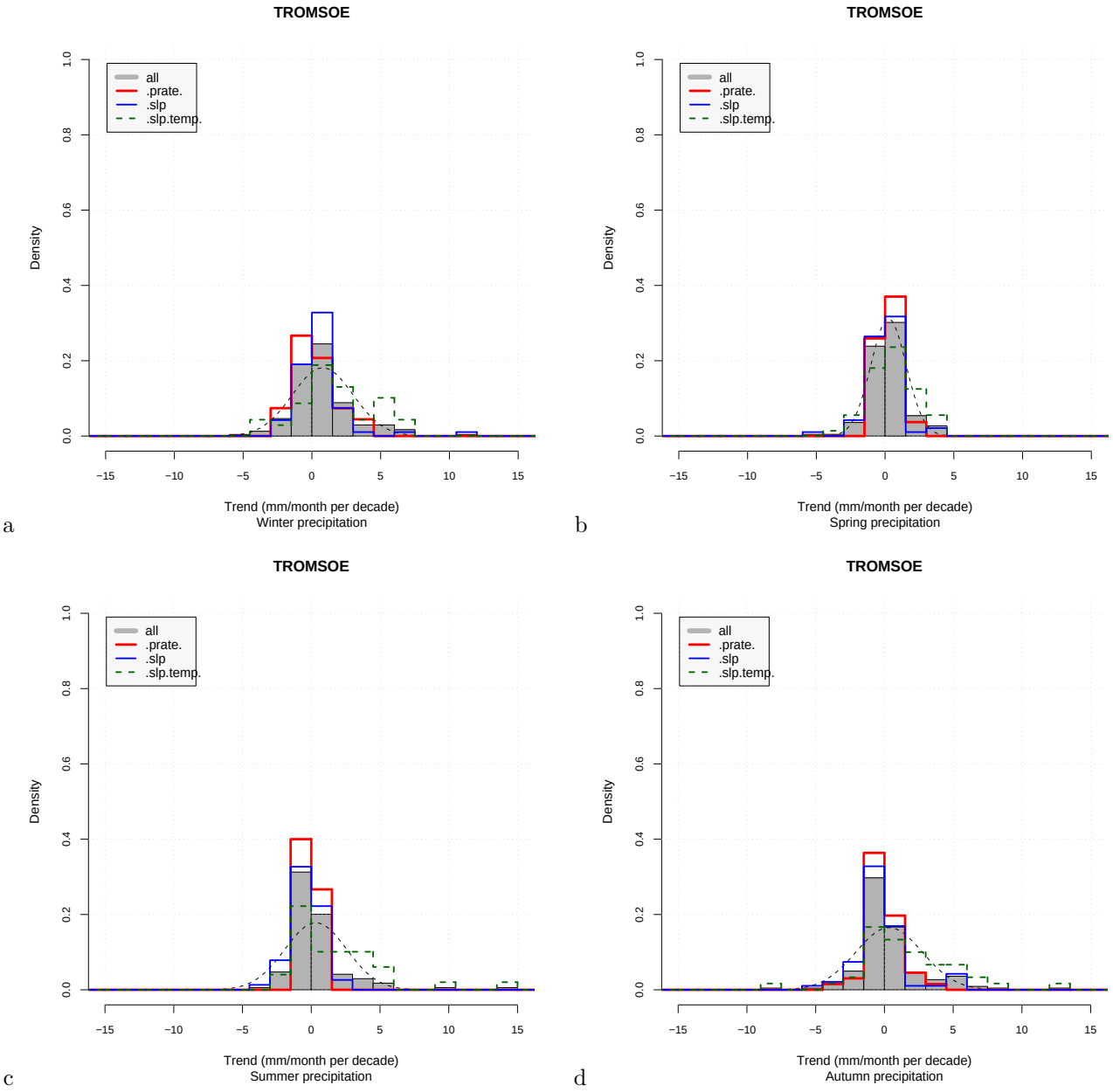

Figure 35. Histograms of downscaled monthly mean precipitation trends for Tromsø presented in Figure 11 for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included.

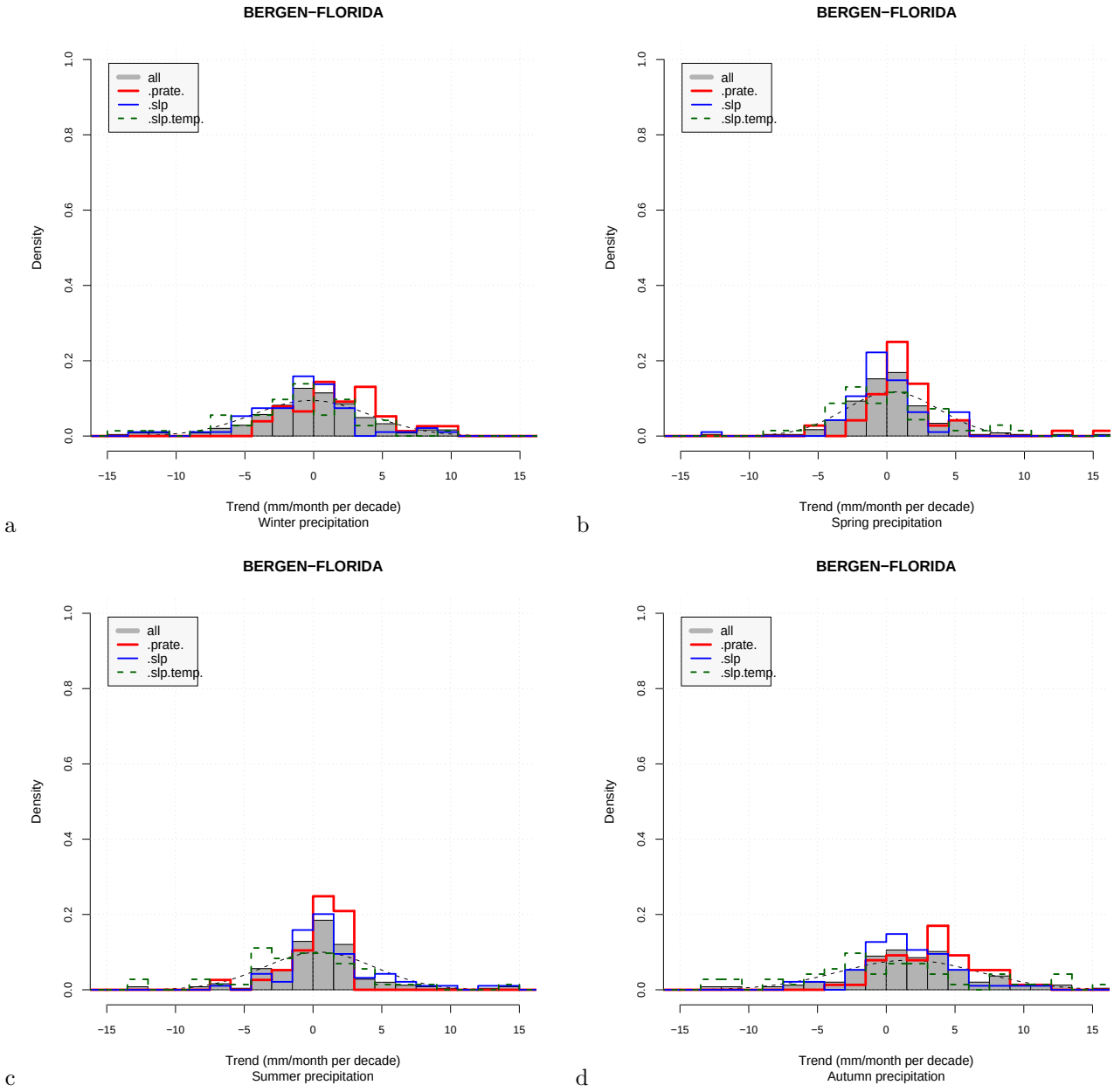

Figure 36. Histograms of downscaled monthly mean precipitation trends for Bergen-Florida presented in Figure 11 for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included.

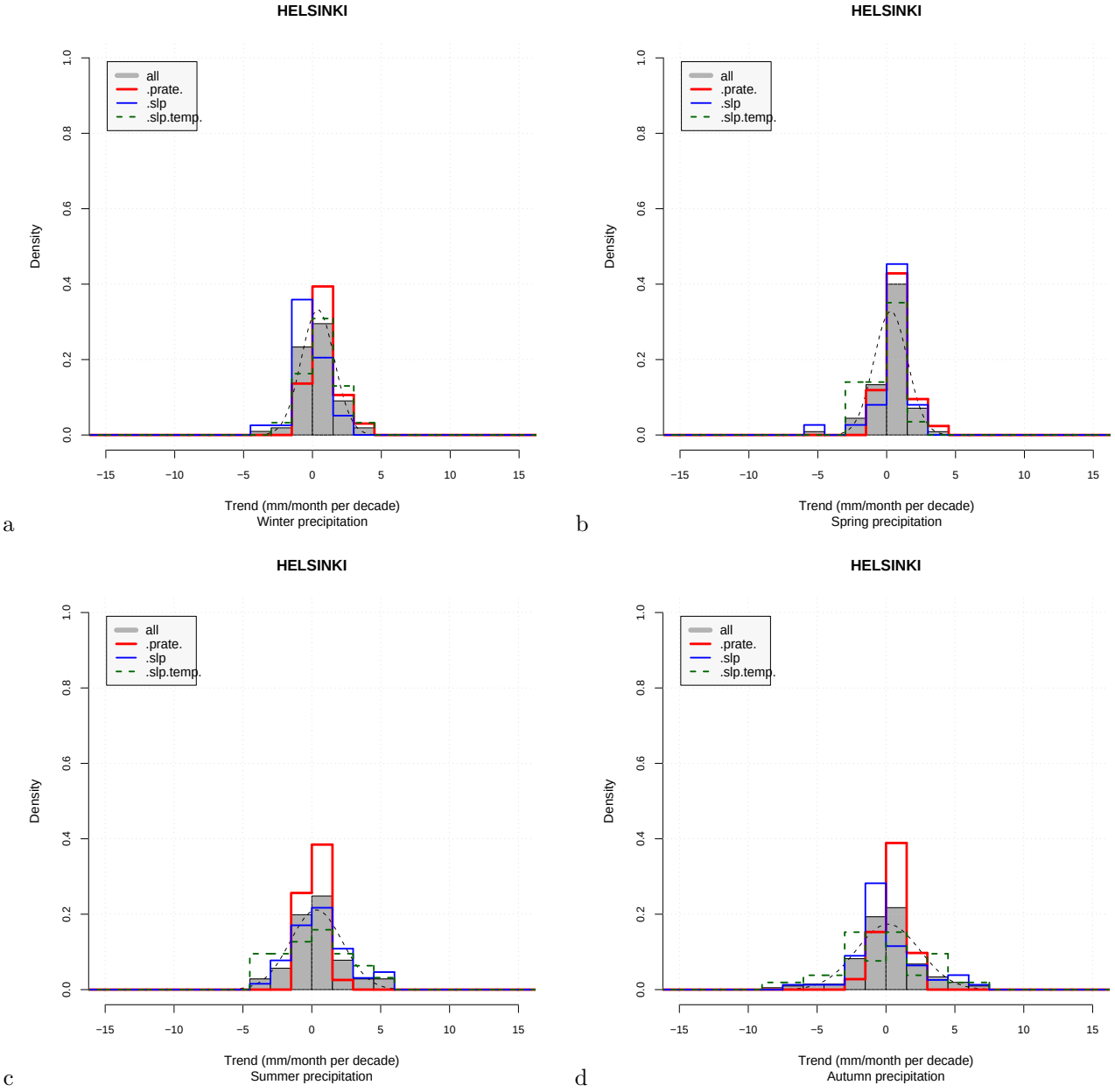

Figure 37. Histograms of downscaled monthly mean precipitation trends for Helsinki presented in Figure 11 for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included.

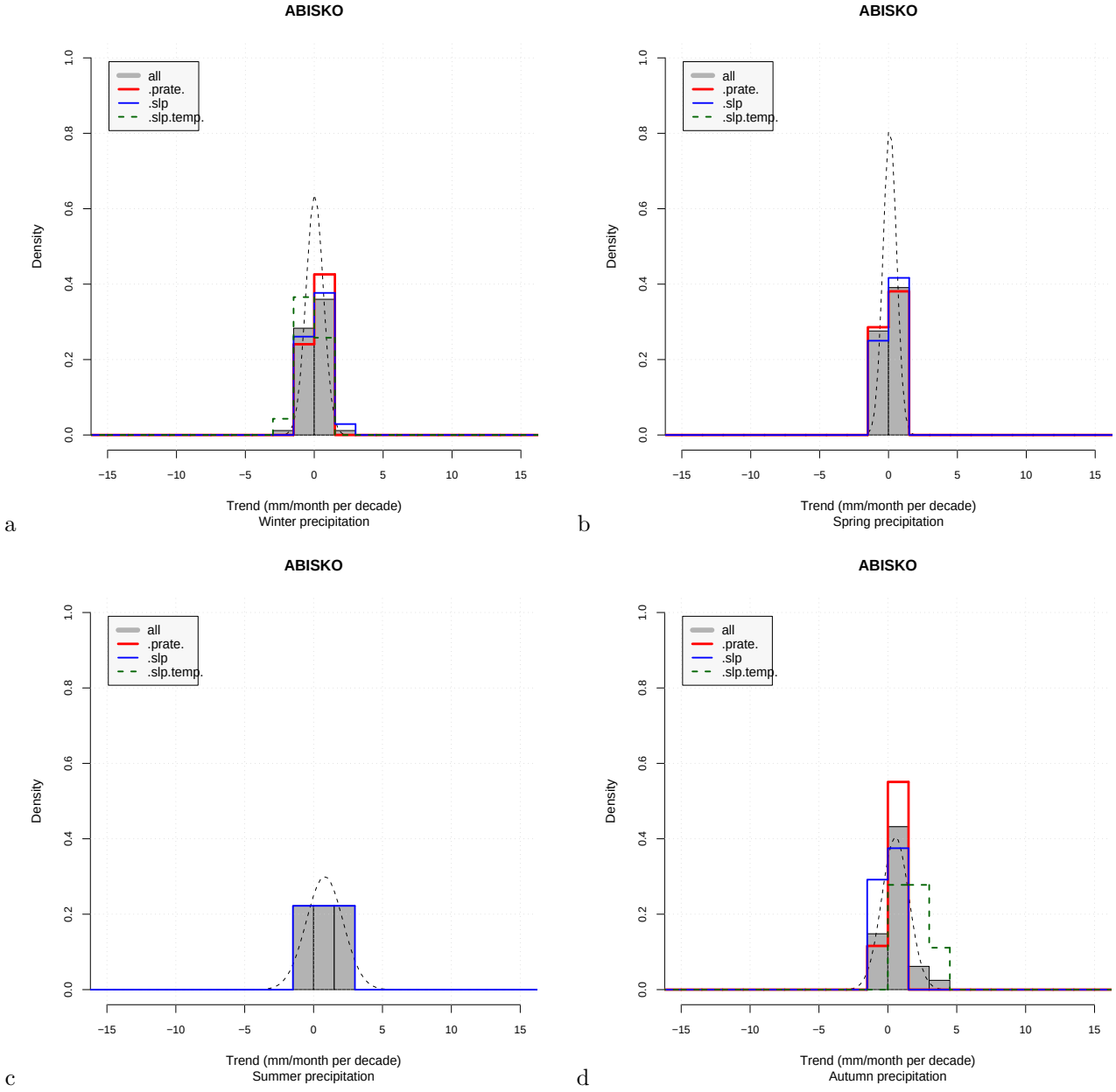

Figure 38. Histograms of downscaled monthly mean precipitation trends for Abisko presented in Figure 11 for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included.

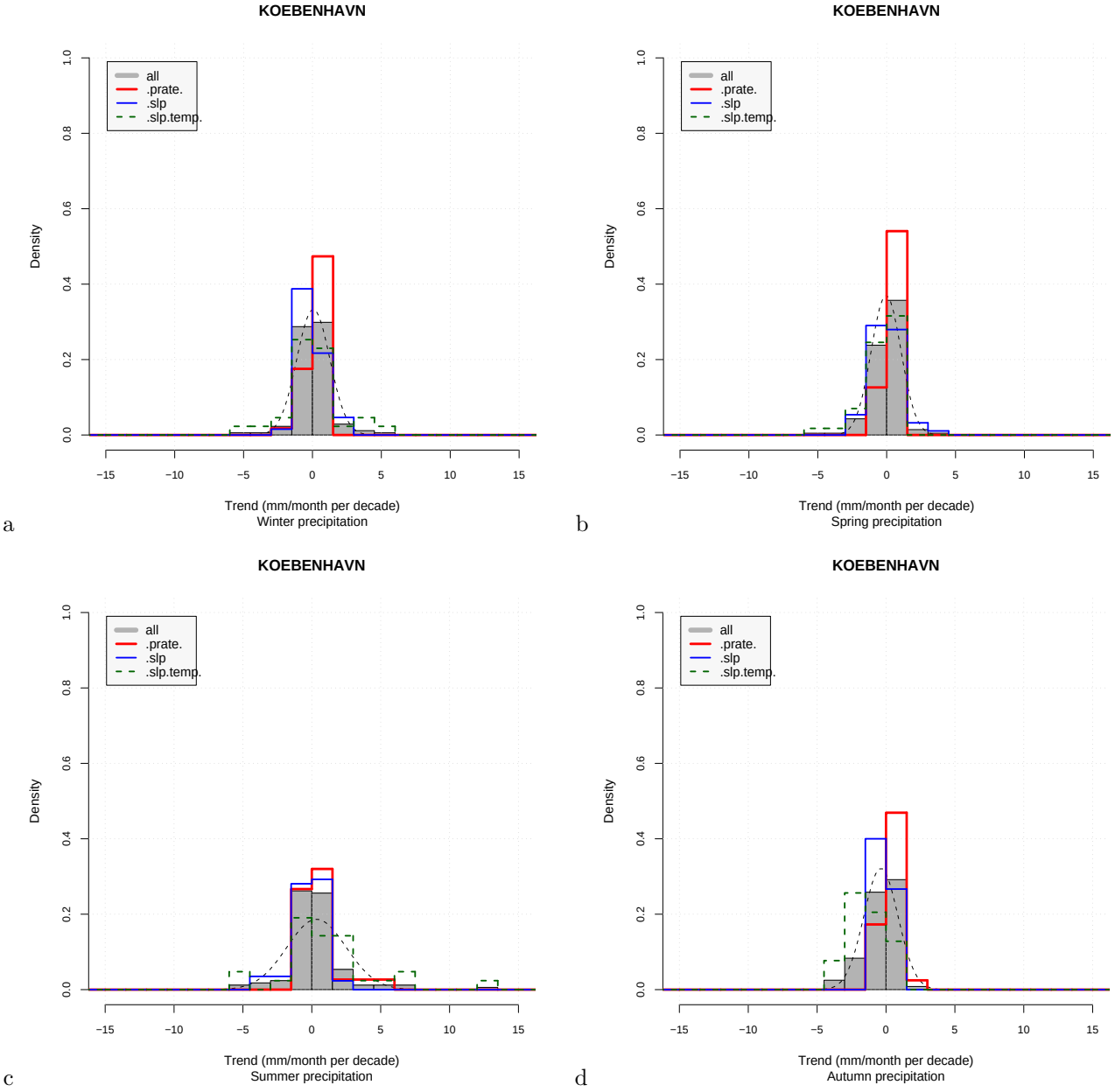

Figure 39. Histograms of downscaled monthly mean precipitation trends for Copenhagen presented in Figure 11 for December–February (a), March–May (b), June–August (c) and September–November (d). The trend estimates were derived using all predictor types and domains, but only R^2 >50% were included.

4 Shortcomings of the downscaling analysis

Figure 40 allows an examination of the downscaled Tromsø summer temperatures in more detail through the inspection of the large-scale temperature anomalies associated with the temperature swings in Tromsø. For realistic scenarios, it is expected to see pronounced weights in the vicinity of the station in Figure 40. The domain choices with most realistic weights are 0°E–50°E/62°N–82°N, 40°W–40°E/52°N–80°N, and 90°W–60°E32°N–80°N. In all these cases, the strongest anomalies are nevertheless located to the east of Tromsø, because the temperature variations are stronger in Finnmark with a more continental type climate than near the coast where they are moderated by the sea (maritime climate). The domain choices 0°E–35°E/57°N–70°N and 20°W–40°E/52°N–70°N do not give weights near the station, and therefore give an indication for less confidence. One problem with these domains is that Tromsø is located at the edge of the predictor domain, and the analysis latches on to the anti-correlation between southern and northern Norway. Benestad (2002b) warned against this type of mistake, using Greenland stations downscaled from large-scale temperature anomalies over Fennoscandia as an illustration.

The downscaled scenarios associated with the domain choices in Figure 40 and the ECHAM4/OPYC3 A2 scenario are presented in Figure 41. Panels (c) and (e) show negative trends whereas (d) and (f) exhibit weak trends. The two domains 0°E–35°E/57°N–70°N and 0°E–50°E/62°N–82°N yield a warming in the ECHAM4/OPYC3 A2 scenario. The most surprising result is the negative trend associated with the 40°W–40°E/52°N–80°N domain. Figure 42 shows how the downscaling with the 40°W–40°E/52°N–80°N domain produces a cooling trend instead of a warming trend: there is an anti-correlation between the temperature on Greenland and Tromsø, and the ECHAM4/OPYC3 A2 integration produces a stronger warming over Greenland than over northern Norway. Furthermore, the EOF pattern has very strong negative weights located over a small Greenland location which does not look very realistic. Benestad et al. (2002) found spurious patterns in their mixed-common EOF products over Greenland, and suggested that the data from this region was of questionable quality. The weights in this region can also be seen in Figure 40(e), and the strong warming trend in this "hot spot" is responsible for the spurious negative trend dervied with this domain choice. Figure 40(f) and Figure 41(f) show the diagnostics for the largest predictor domain 90°W–60°E32°N–80°N, and although this analysis includes Greenland, the trend produced for the July temperature in Tromsø is still positive, albeit weak. Figure 43 shows the same dignostics as Figure 42, but for the 90°W–60°E32°N–80°N domain. These diagnostics suggest that the leading common EOF represents a warming over the continental as well as over Greenland, and the east-west dipole associated with the NAO is less pronounced with the larger domain. Too extensive domain are dangerous, and it is better with too small than too large although it is desirable to have a predictor domain that is larger than the skillfull spatial scale Grotch & MacCracken (1991).

5 Best estimates and comparison with earlier work

The bulk statistics presented above has revealed that using SLP only as predictor does not capture the temperature or precipitation trends, and it is therefore necessary to remove the cases for which SLP has been the sole predictor in order to get realistic trend statistics. The subsequent analysis has also dropped the results from CCCma and the CSIRO models. The above discussion has also focussed on the SRES B2 scenario, but it is desirable to look at the results from more than one emission scenario. Below, the rate of change (in °C/decade or [mm/month]/decade) derived for the 2000-2100 interval using the A2 and B2 emission are furthermore presented together with trend rates for 1980–2050 from the IS92a scenarios reported in Benestad (2002a). Figure 44 presents histograms of the trend rates for January (a), April (b), July (c) and October (d). The ensembles for the older IS92a scenarios are larger (48 for temperature; 50 for precipitation, representing 17 independent GCM integrations) than the more recent SRES scenarios (ensemble size is 26, but representing only 6 independent GCM integrations). Hence, the comparison in Figure 44 only gives a crude description of the results. For the temperature trends in Oslo, the new scenarios have similar distributions to those based on the IS92a for January and October, but indicate stronger warming in April and July. The B2 scenario gives slightly weaker winter warming in Oslo whereas the A2 scenario indicates similar warming rate. In the autumn, the B2 indicates slightly slower warming while the A2 produces a somewhat stronger warming. In spring the A2 and B2 give similar warming rates.

The histograms for precipitation rates are "noisy" and not well-defined (Figure 45). However, the

greatest differences between the various scenarios are seen in July. The new SRES-based scenarios indicate drier future July months in Oslo.

The results for the temperature in Bergen (Figure 46) suggest some difference between the various emission scenarios for Bergen, and the more recent SRES-based scenarios give slightly wider distributions. The B2-based derivations point to weaker warming in January and similar trends in October, but stronger warming for the other months. The A" scenario produces stronger warming for all the cases. The histograms for rainfall in Figure 47 are difficult to interpret, as there is a tendency of several peaks interspersed with low counts. Because of a substantial scatter in the estimates, the ensemble size for the precipitation is probably too small for a trustworthy interpretation. The new results are nevertheless roughly similar to the old estimates based on the IS92a emission scenario (Benestad, 2002a).

Figure 48 shows the histograms for the temperature trends downscaled for Tromsø. Apart for in April, there are small differences between the various scenarios. The portion of negative July temperature trends derived from the SRES A2 scenario is due to shortcomings mentioned above, and the confidence in these results need to be questioned (see Figures 41–43 and section 4). Tromsø is situated near the edge of the predictor domain used here. Figure 49 shows the histograms for the downscaled precipitation trends in Tromsø. A number of outliers are visible in the SRES scenarios for all months. There is a slight hint of an increase in the January precipitation relative to the older IS92a-based results (a), however, the distribution for April is largely similar for the different scenarios (b). One difference is that the A2- and B2-based scenarios yield greater scatter. For July, the SRES scenarios produce much broader distributions (c), and in October, the A2 scenarios suggests marginally drier conditions than the B2 and the older IS92a scenarios.

6 Discussion & Conclusions

Mixed-common EOFs were examined in order to evaluate the output of the various climate models. The ECHAM4/OPYC3 and the HadCM3 model exhibited the most realistic spatial structures and magnitudes. The CCCma and the CSIRO GCMs, on the other hand, did not produce realistically looking spatial patterns of variability, possibly due to their low spatial resolution. There were also some indications suggesting discrepancies between the spatial modes in the NCAR-PCM results and the observations, possibly due to unrealistically cold conditions in the vicinity of Greenland and pronounded variations in the vicinity of the ice-edge. SLP by itself is not appropriate for downscaling temperature as associated temperature trends exhibited no clear warming. SLP doesn't capture trends in atmospheric moisture, and may underestimate the precipitation as well. New empirical models based on predictors of large-scale precipitation rate appear more promising than the more traditional SLP-based derivations. Predictors consisting of mixed temperature-SLP fields and temperature-only fields produced similar skill scores and trend estimates for temperature.

The downscaled scenarios confirmed the general tendency of warming. There was a small number of cases where the downscaled temperature trends were negative, and the majority of these involved the Arctic* stations as well as the less realistic GCMs CCCma and CSIRO. The GCMs with higher confidence did nevertheless also produce a few cases with apparent cooling trends. Since these were not credible, it was speculated as to whether these results were due to inappropriate predictor domains. This point raises the question of whether an objective scheme can be deviced for choosing an optimal predictor domain. Future work will address this question by using spatial correlation maps as a basis for selecting the predictor area. Large inter-annual variations in the high-latitudes may also have affected the trend estimates.

The inspection of interpolated values for the annual cycle in precipitation may at first sight suggest that the description of precipitation by GCMs is problematic: None of the GCMs did skillfully reproduce the phase, amplitude of precipitation measured at the stations or interpolated by the NCEP re-analysis. This underlines the danger of using interpolated precipitation values from the GCMs. By applying empirical downscaling, it can be demonstrated that description of local precipitation is on the other hand realistic. The discrepancies associated with the seasonally varying rainfall in some locations raises questions regarding the confidence in long-term changes in precipitation. It is argued that for these sites local precipitation is only weakly affected by changes in the boundary conditions and more strongly

*Here referring to Svalbard, Bjrnya, and Jan Mayen

affected by internal chaotic dynamics and local cloud formation. The differences between the interpolated results also may have been due to local geographical influence. A comparison of the seasonal rainfall variation in Bergen, on the other hand, suggests a good correspondence between the GCMs and the observations. Furthermore, much of the summer rainfall is due to local small-scale convection rather than weather systems extending over 100km. The assessment of the downscaling models based on large-scale precipitation as a predictor points to considerable skill for anomalies. The striking inter-model similarity in the EOF patterns and similar magnitude in the re-analysis and GCM output suggests that the GCMs reproduce the variability of large-scale precipitation quite well over the oceans.

The precipitation in northern Europe is strongly influenced by the type of circulation pattern, and there are large year-to-year variations in the rainfall. There is no clear trend in circulation patterns important for the local temperatures when the SLP is one of the major predictors, resulting in no clear trend in precipitation. Trends derived from large-scale precipitation rate only, on the other hand, point to slight shifts in the trend distribution about zero. The conclusions for the downscaled precipitation scenarios from this study is similar to findings reported in Benestad (2002a). It is also important to note that the results derived in that study, although also employing the common EOF frame work, were based on a different linear model using Canonical Correlation Analysis instead of regression and was coded for Matlab as opposed `R`. The gridded observations used also differed since they were based on Benestad (2000a) whereas the derivation of these results involved the NCEP reanalysis (the two gridded data sets are not entirely independent of each other). The similarity between the results despite these differences therefore point to a degree of robustness in these results. Furthermore, the lack of clear trends in precipitation is also consistents with findings by Ruosteenoja et al. (2003) who noted that simulated precipitation changes on sub-continental scales, considered on their own, are predominantly not statistically significant to modelled natural variability. They also obtain both positive and negative trends for one given region, possibly due to the finding that several models simulate precipitation changes of opposite sign with various forcing scenarios and low signal-to-noise ratios.

Although there is a fairly robust picture for the temperature with a general warming trend, the results for precipitation exhibit a great scatter. Hence, there is a need for larger ensembles in order to produce robust statistics that would be required to resolve the issue of trend for these locations. Large ensembles may become available in the near future from the `climatepredction.net` project (Allen, 2003).

Acknowledgments:This work was done under the Norwegian Regional Climate Development under Global Warming (RegClim) programme, and was supported by the Norwegian Research Council (Contract NRC-No. 120656/720) and the Norwegian Meteorological Institute. The analysis was carried out using the `R` (Ellner, 2001; Gentleman & Ihaka, 2000) data processing and analysis language, which is freely available over the Internet (URL http://www.R-project.org/).

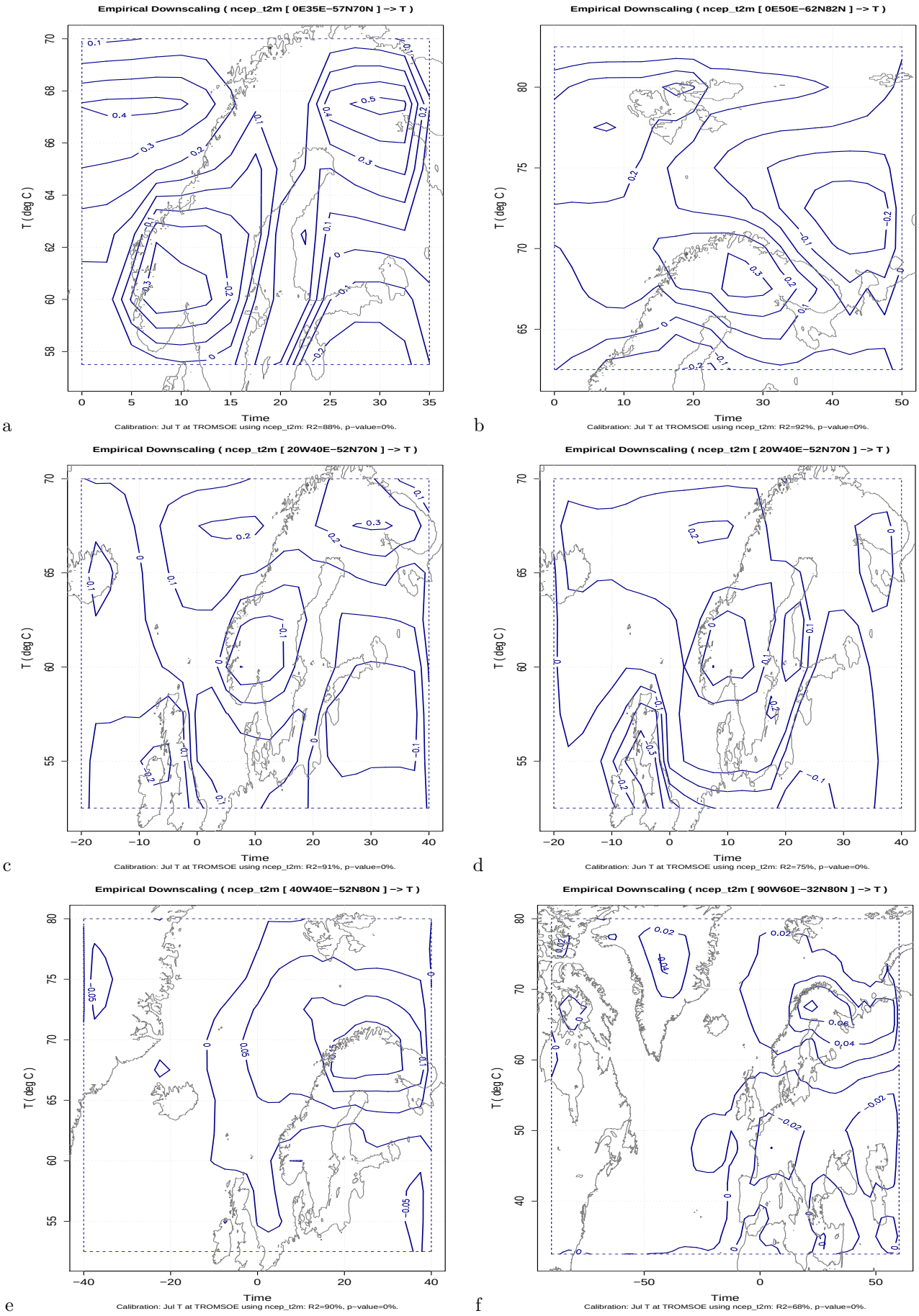

Figure 40. Spatial pattern of the large-scale temperature anomaly that is associated with July (June in Panel d) temperature variations in Tromsø. The different panels show the patterns obtained using different domain choices.

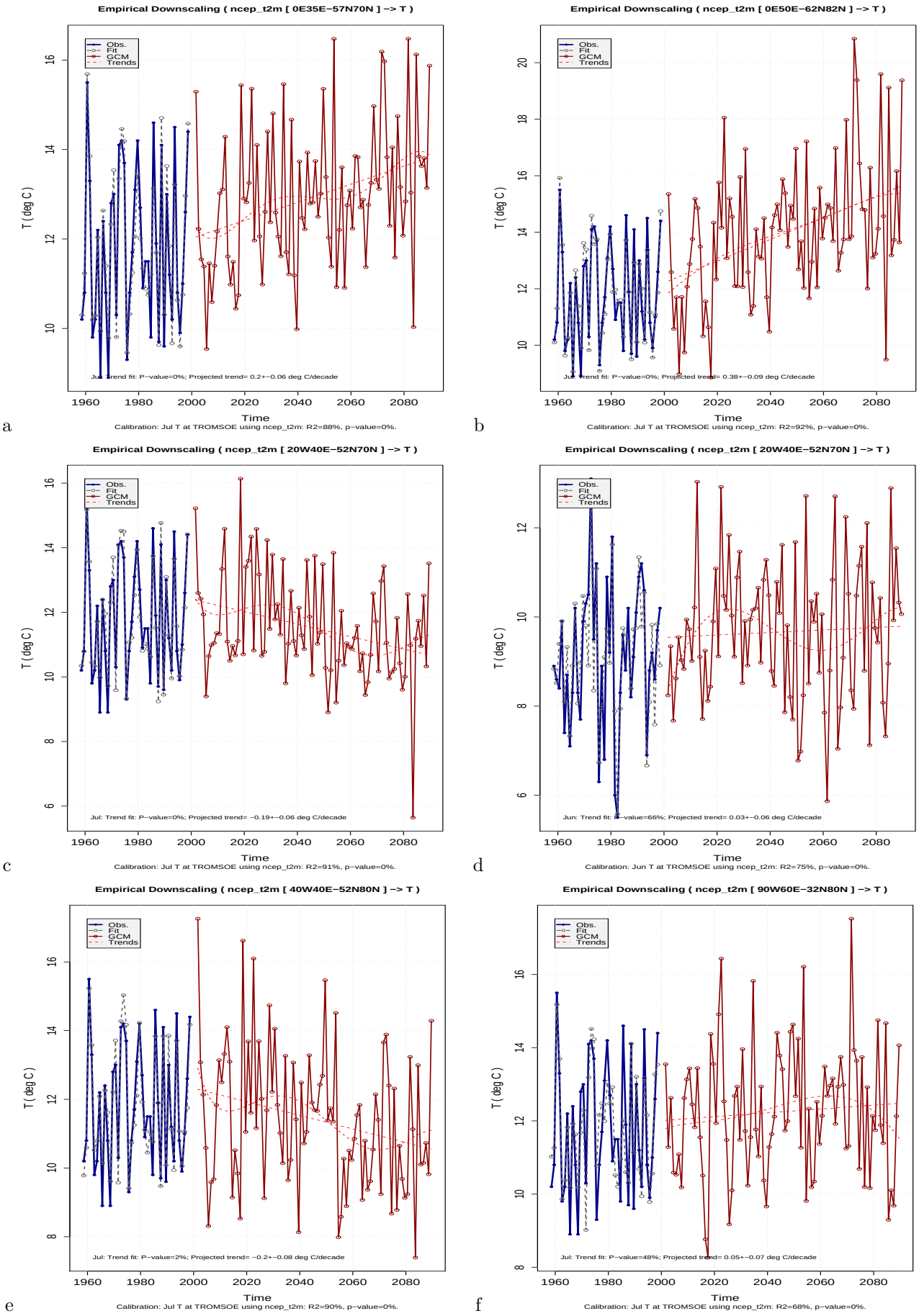

Figure 41. Downscaled scenarios for July (June in Panel d) temperature variations in Tromsø associated with the large-scale temperature structures in Figure 40. ECHAM4/OPYC3 A2 scenario.

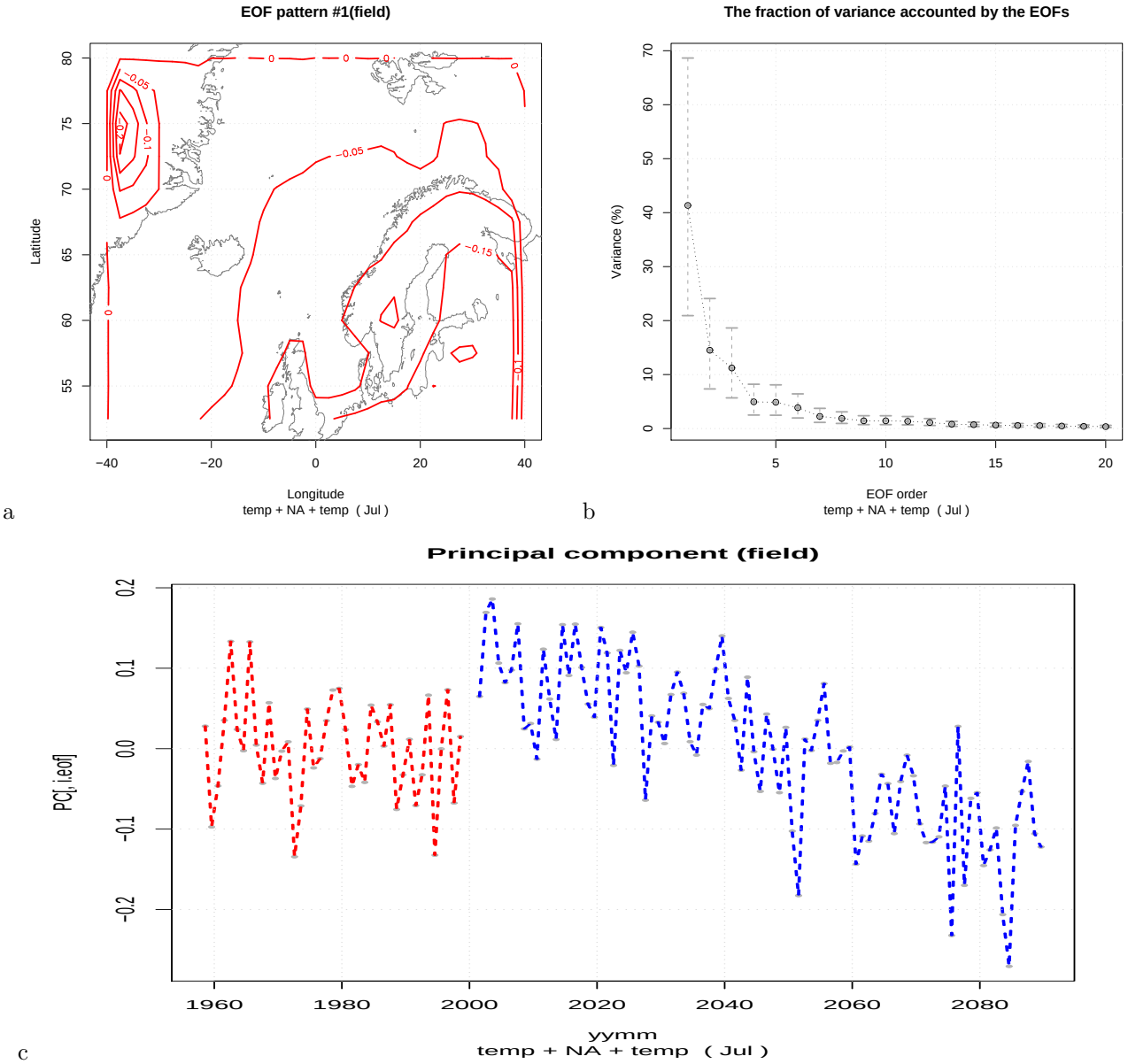

Figure 42. common EOF products for the ECHAM4/OPYC3 A2 July temperature.

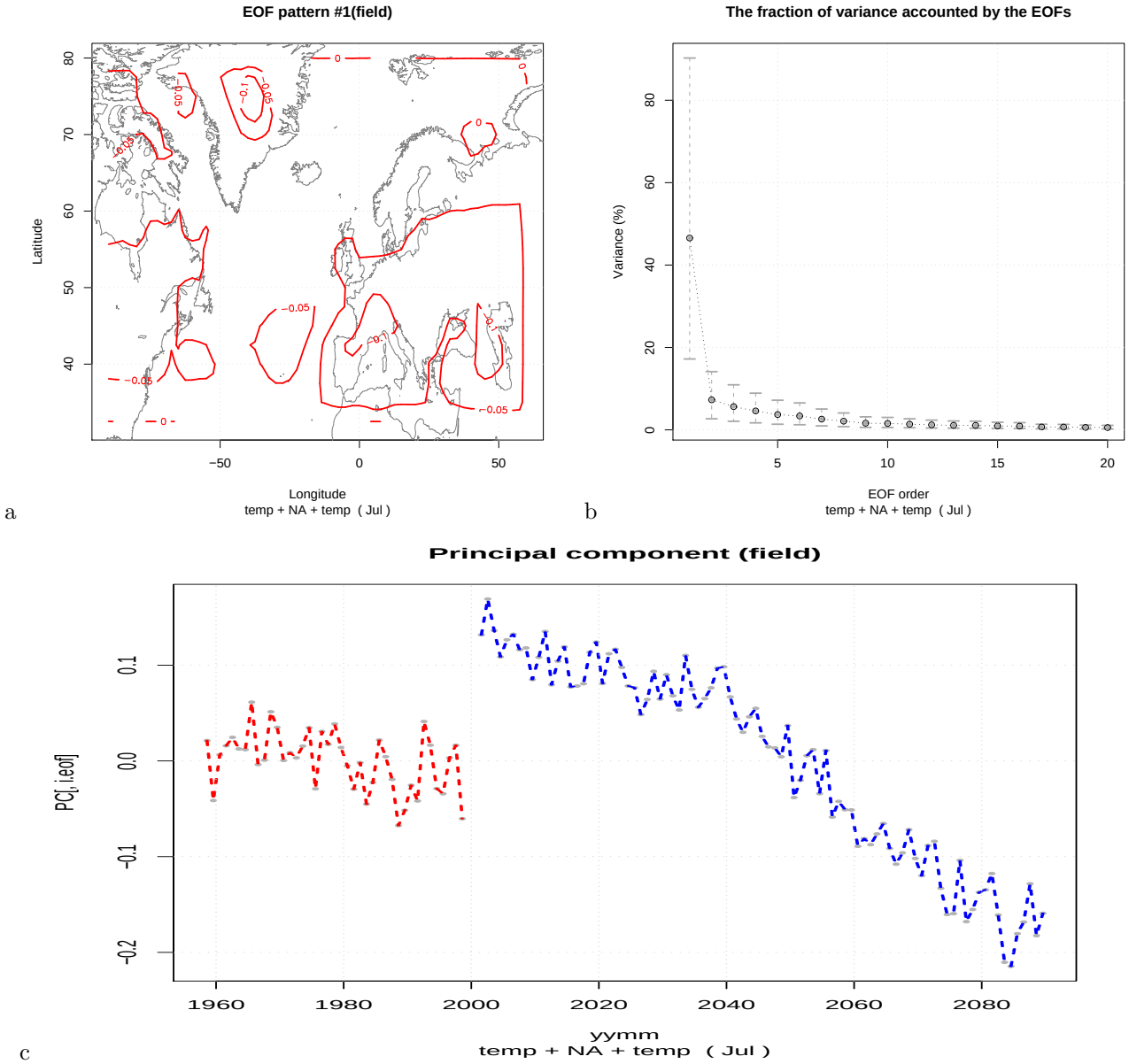

Figure 43. common EOF products for the ECHAM4/OPYC3 A2 July temperature.

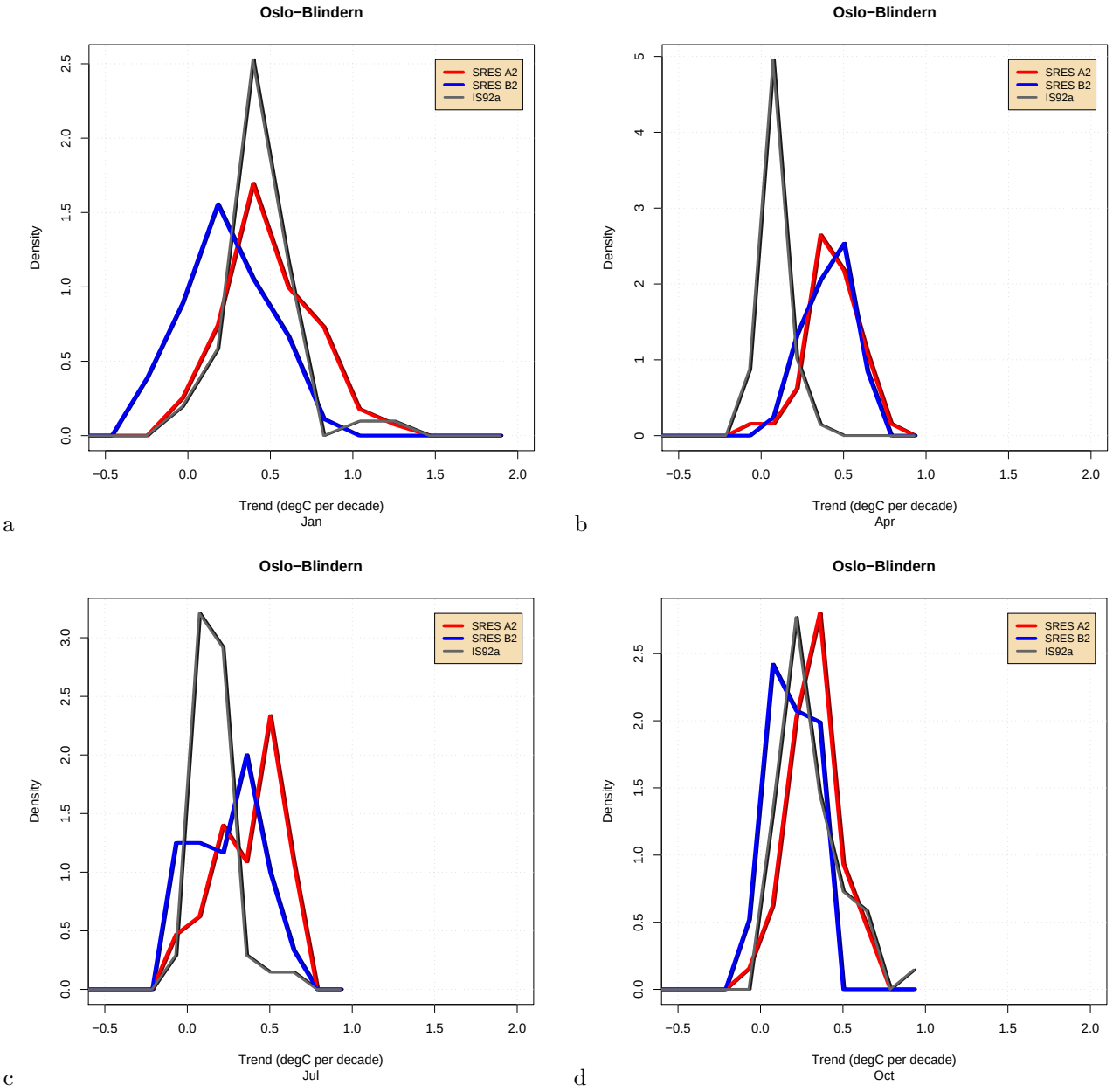

Figure 44. Histograms of downscaled monthly mean temperature trends for Oslo-Blindern for January (a), April (b), July (c) and October (d). The trend estimates were derived using the 'temp.' and '.slp.temp.' predictor types only, all domains, and only the results for which R^2 >50% were included.

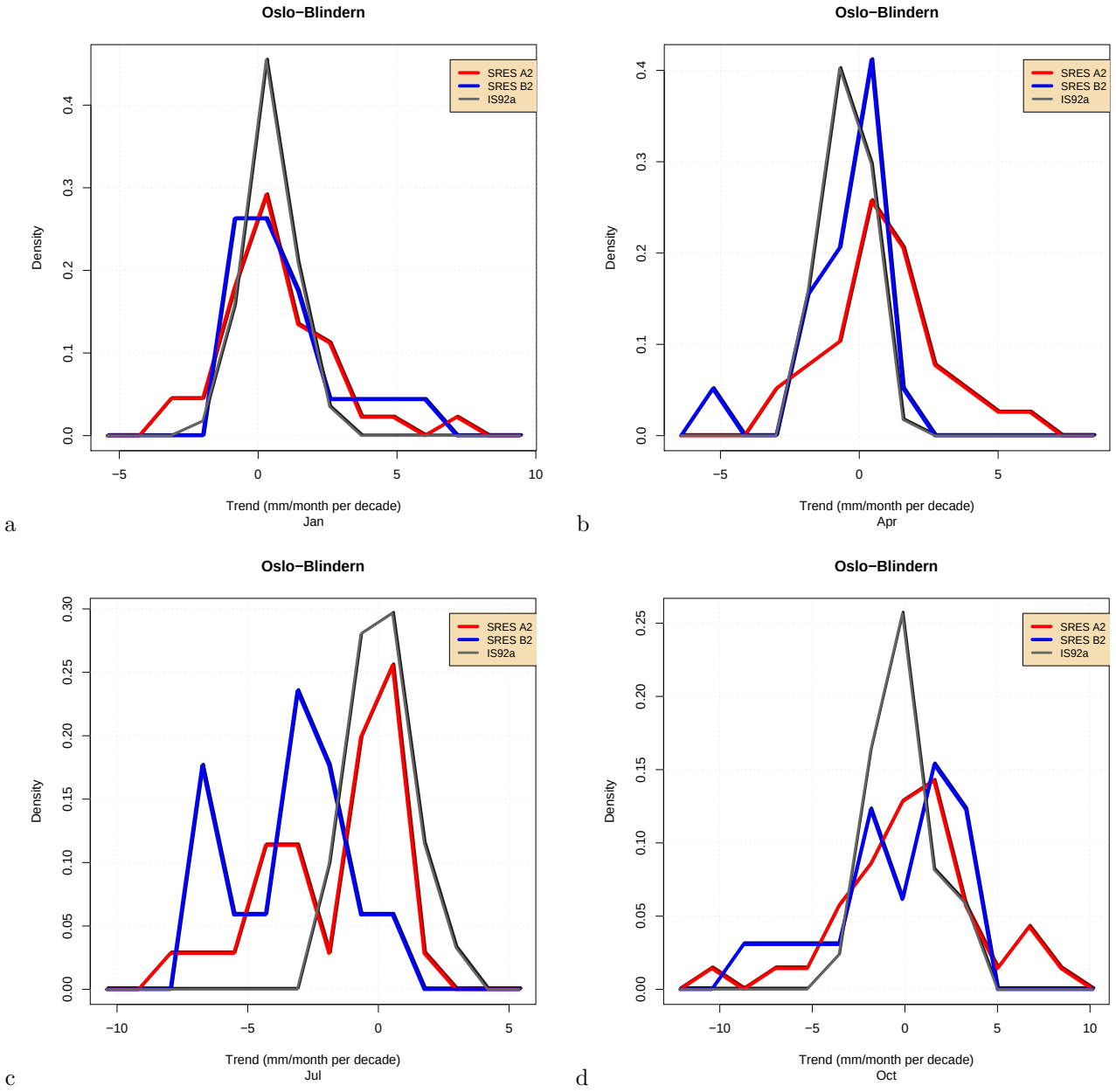

Figure 45. Histograms of downscaled monthly mean precipitation trends for Oslo-Blindern for January (a), April (b), July (c) and October (d). The trend estimates were derived using the 'temp.' and '.slp.temp.' predictor types only, all domains, and only the results for which R^2 >50% were included.

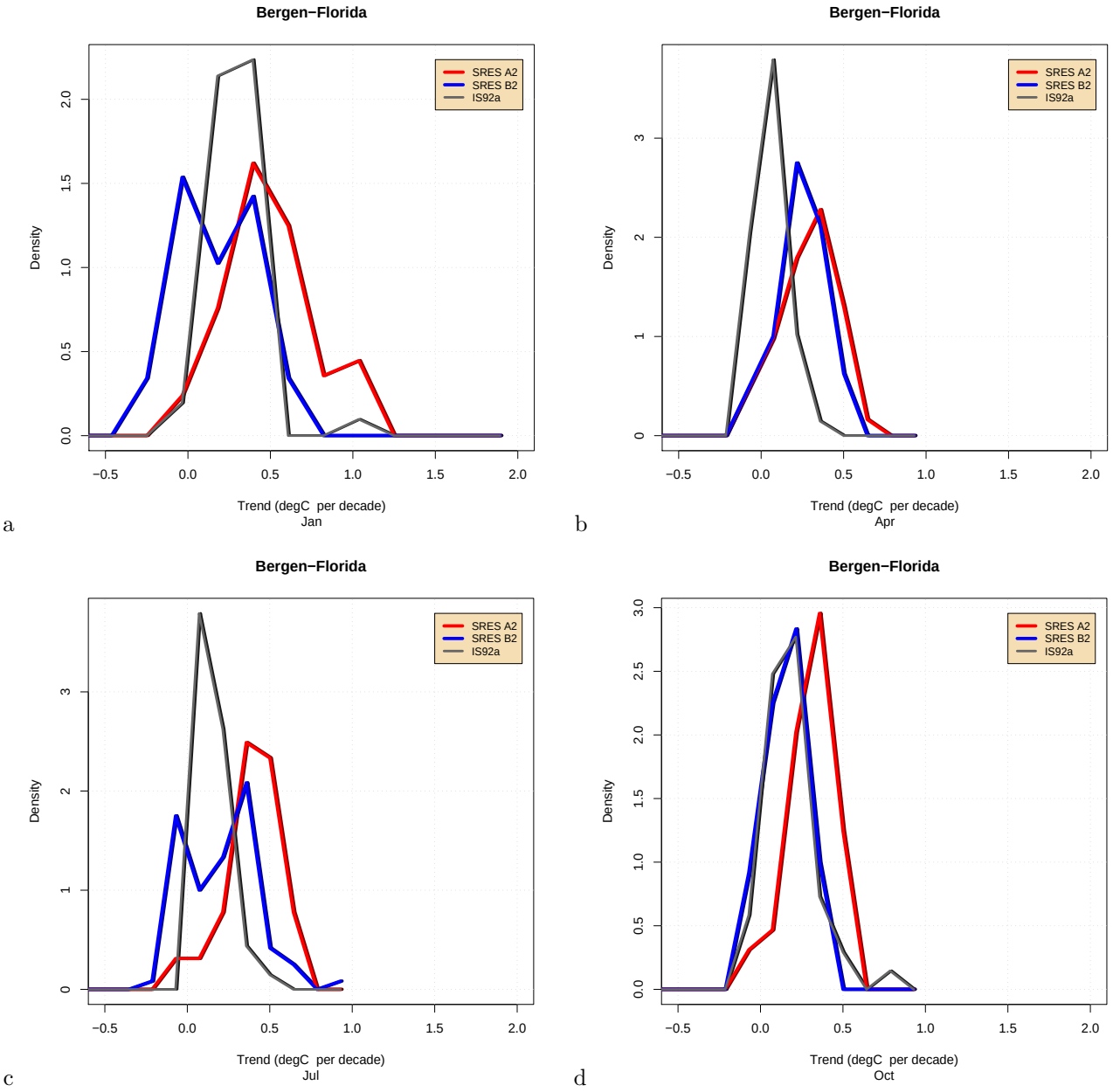

Figure 46. Histograms of downscaled monthly mean temperature trends for Bergen-Florida for January (a), April (b), July (c) and October (d). The trend estimates were derived using the 'temp.' and '.slp.temp.' predictor types only, all domains, and only the results for which R^2 >50% were included.

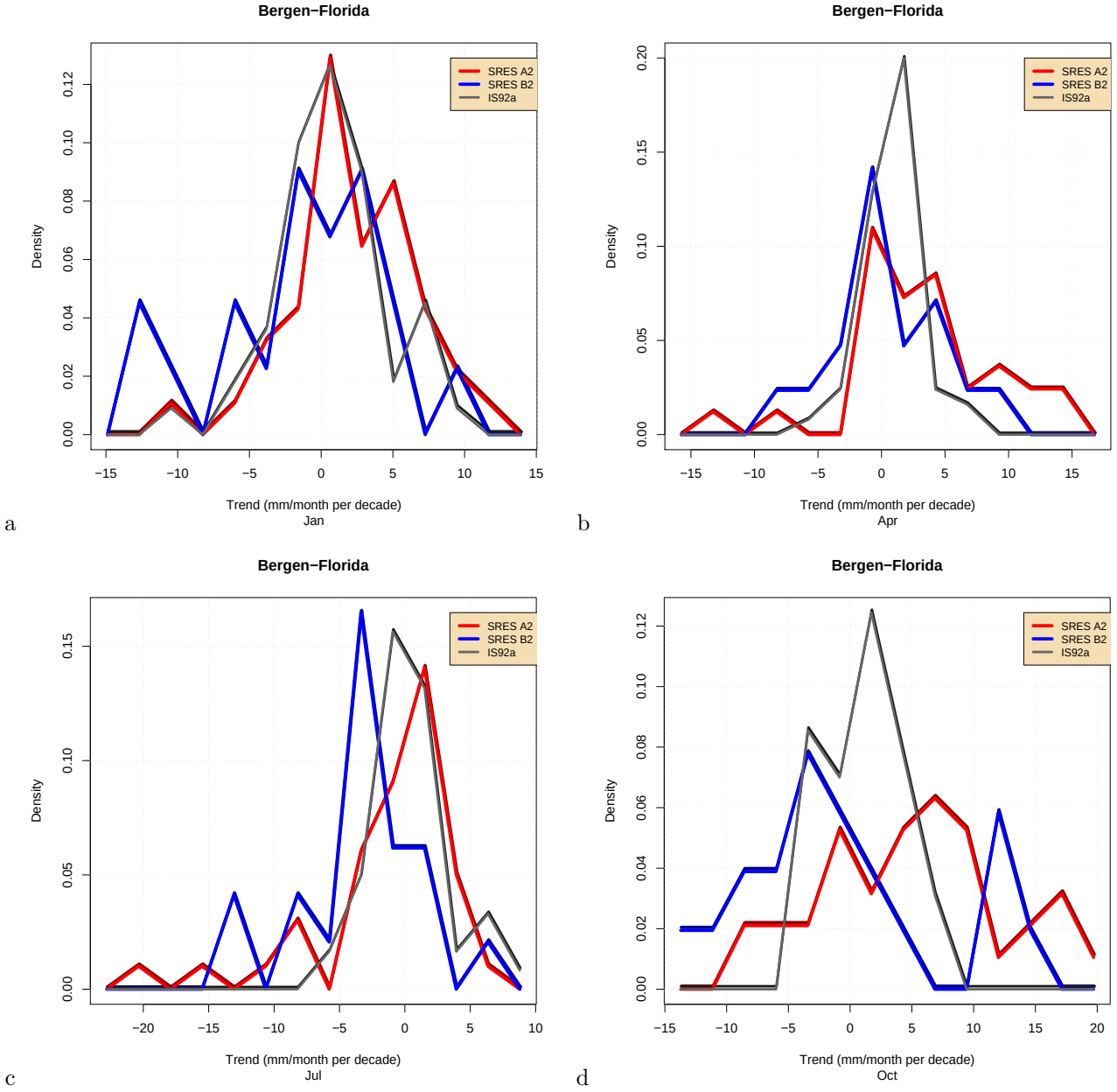

Figure 47. Histograms of downscaled monthly mean precipitation trends for Bergen-Florida for January (a), April (b), July (c) and October (d). The trend estimates were derived using the 'temp.' and '.slp.temp.' predictor types only, all domains, and only the results for which R^2 >50% were included.

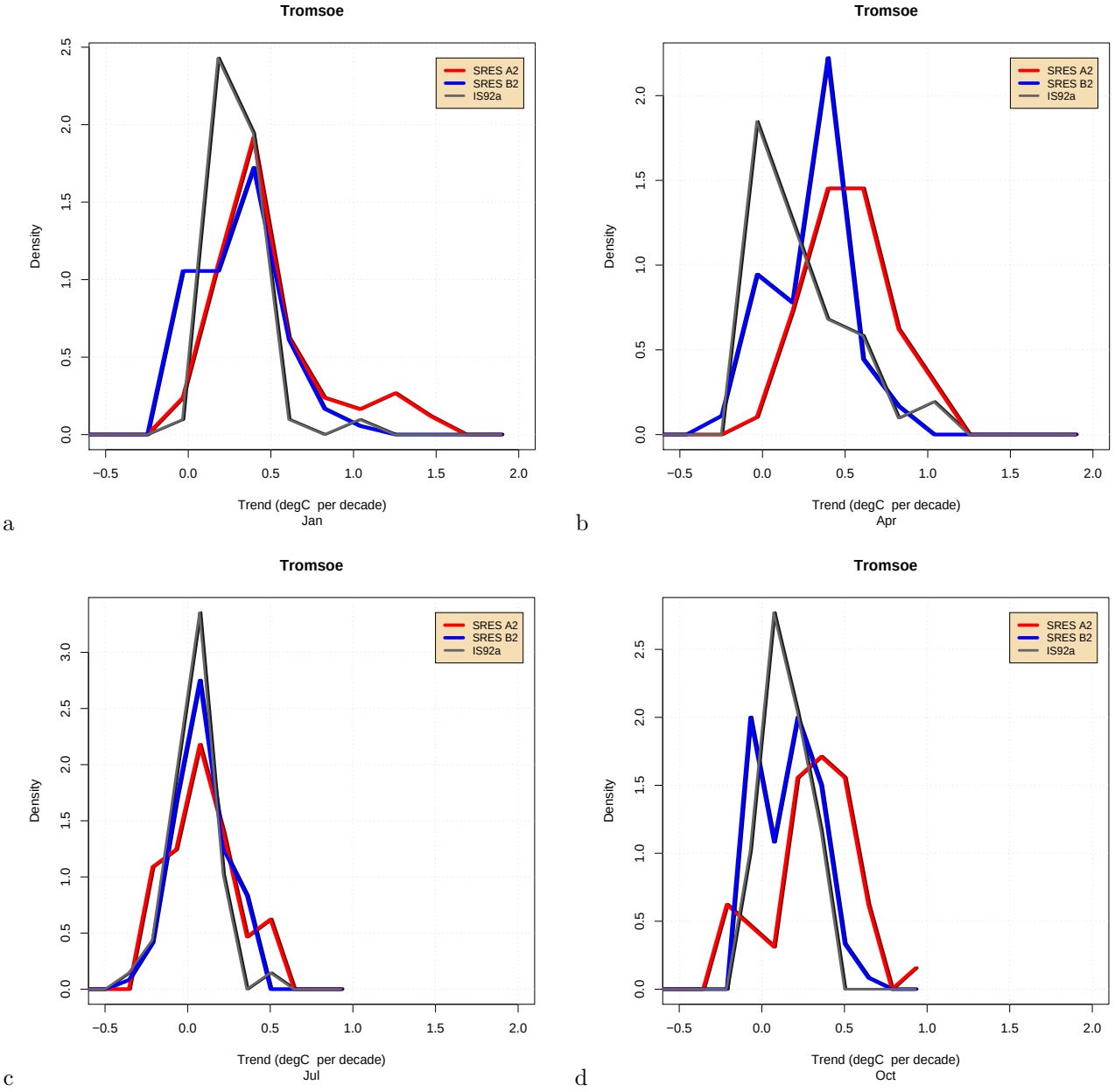

Figure 48. Histograms of downscaled monthly mean temperature trends for Tromsø for January (a), April (b), July (c) and October (d). The trend estimates were derived using the 'temp.' and '.slp.temp.' predictor types only, all domains, and only the results for which R^2 >50% were included.

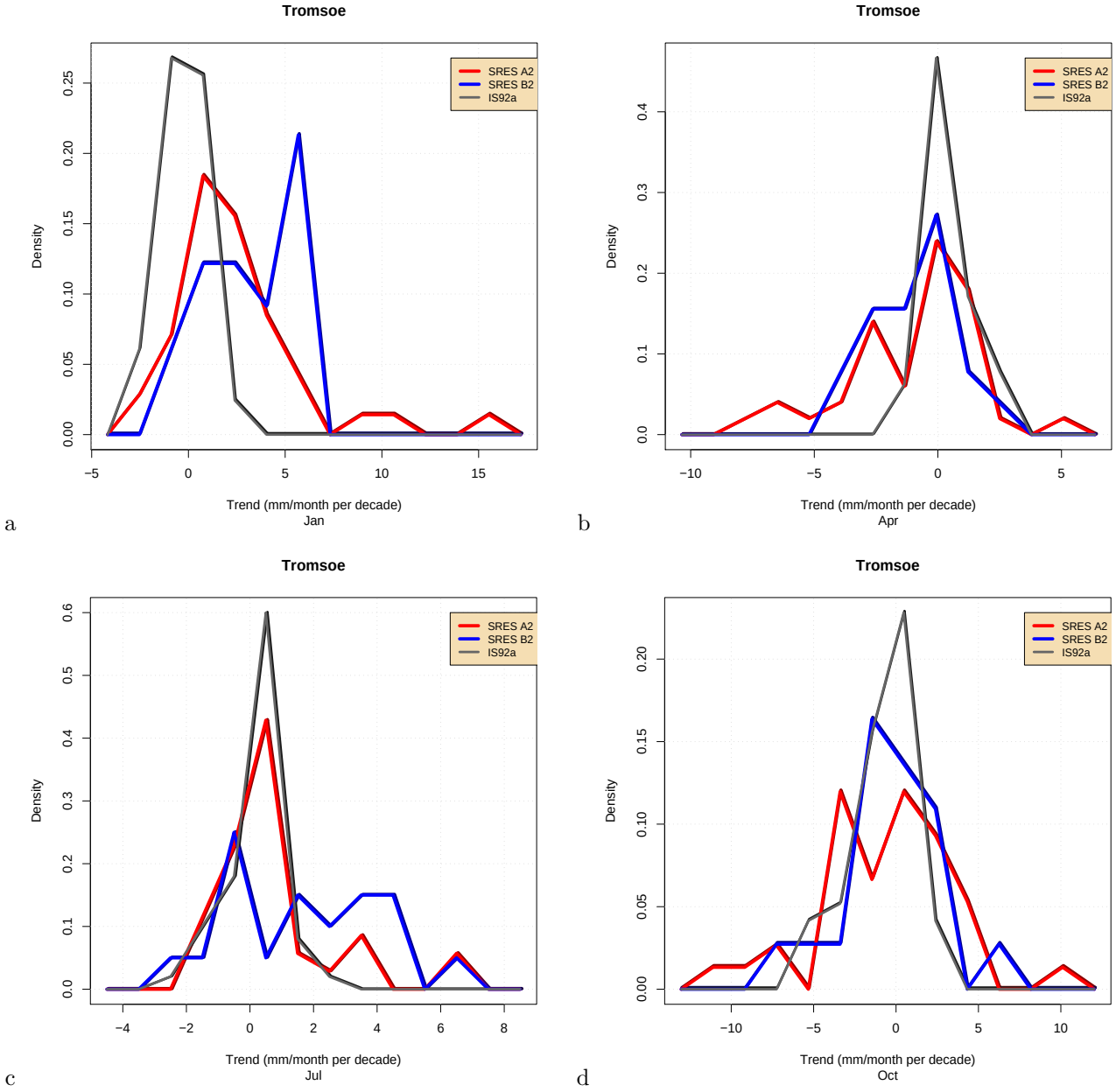

Figure 49. Histograms of downscaled monthly mean precipitation trends for Tromsø for January (a), April (b), July (c) and October (d). The trend estimates were derived using the 'temp.' and '.slp.temp.' predictor types only, all domains, and only the results for which R^2 >50% were included.

References

Addison, P., 2004. The little wave with the big future. Physics World, March, 35–39.

Allen, M.R., 2003. Possible or probable? Nature, 425, 242.

Barnett, T.P., 1999. Comparison of Near-Surface Air Temperature Variability in 11 Coupled Global Climate Models. Journal of Climate, 12, 511–518.

Benestad, R.E., 2000a. Analysis of gridded sea level pressure and 2-meter temperature for 1873-1998 based on UEA and NCEP re-analysis II. KLIMA 03/00. DNMI, PO Box 43 Blindern, 0313 Oslo, Norway.

Benestad, R.E., 2000b. Future Climate Scenarios for Norway based on linear empirical downscaling and inferred directly from AOGCM results. KLIMA 23/00. DNMI, PO Box 43 Blindern, 0313 Oslo, Norway.

Benestad, R.E., 2001a. The cause of warming over Norway in the ECHAM4/OPYC3 GHG integration. International Journal of Climatology, 21, 371–387.

Benestad, R.E., 2001b. A comparison between two empirical downscaling strategies. Int. J. Climatology, 21, 1645–1668. DOI 10.1002/joc.703.

Benestad, R.E., 2002a. Empirically downscaled multi-model ensemble temperature and precipitation scenarios for Norway. Journal of Climate, 15, 3008–3027.

Benestad, R.E., 2002b. Empirically downscaled temperature scenarios for northern Europe based on a multi-model ensemble. Climate Research, 21, 105–125.

Benestad, R.E., 2003a. clim.pact-V.1.0. KLIMA 04/03. The Norwegian Meteorological Institute, PO Box 43 Blindern, 0313 Oslo, Norway (www.met.no).

Benestad, R.E., 2003b. Downscaling analysis for daily and monthly values using clim.pact-V.0.9. KLIMA 01/03. met.no, PO Box 43 Blindern, 0313 Oslo, Norway (www.met.no).

Benestad, R.E., 2003c. A first-order evaluation of climate outlooks based on the IPCC A2 and B2 SRES emission scenarios. KLIMA 03/03. The Norwegian Meteorological Institute, PO Box 43 Blindern, 0313 Oslo, Norway (www.met.no).

Benestad, R.E., 2004. Tentative probabilistic temperature scenarios for northern Europe. Tellus, 56A, 89–101.

Benestad, R.E., Hanssen-Bauer, I., & Førland, E.J., 2002. Empirically downscaled temperature scenarios for Svalbard. Atmospheric Science Letters, September 18, doi.10.1006/asle.2002.0051.

Bretherton, C.S, Smith, C., & Wallace, J.M., 1992. An Intercomparison of Methods for finding Coupled Patterns in Climate Data. Journal of Climate, 5, 541–560.

Ellner, S.P., 2001. Review of R, Version 1.1.1. Bulletin of the Ecological Society of America, 82(April), 127–128.

Flury, B., 1988. Common Principal Components and Related Multivariate Models. Wiley Series in Probability and Mathematical Statistics. New York: Wiley.

Gentleman, R., & Ihaka, R., 2000. Lexical Scope and Statistical Computing. Journal of Computational and Graphical Statistics, 9, 491–508.

Goodess, C., Osborn, T., & Hulme, M., 2003 (April). The identification and evaluation of suitable scenario development methods for the estimation of future probabilities of extreme weather events. Technical Report 4. Tyndall Centre, School of Environemtal Sciences, Univ. East Anglia, Norwich.

Grotch, S., & MacCracken, M., 1991. The use of general circulation models to predict regional climate change. Journal of Climate, 4, 286–303.

Hanssen-Bauer, I., Førland, E.J., Haugen, J.E., & Tveito, O.E., 2003. Temperature and precipitation scenarios for Norway: Comparison of results from dynamical and empirical downscaling. Climate Research, 25, 15–27.

Imbert, A., & Benestad, R.E., 2003. An improvement of analog model strategy for more reliable local climate change scenarios. Theoretical and Applied Climatology, submitted.

IPCC., 2001. Climate Change 2001: The Scientific Basis. Summary for Policymakers. WMO.

North, G.R., Bell, T.L., & Cahalan, R.F., 1982. Sampling Errors in the Estimation of Empirical Orthogonal Functions. Monthly Weather Review, 110, 699–706.

Ruosteenoja, K., Carter, T.R., Jylhä, K., & Tuomenvirta, H., 2003. Future climate in world regions: an intercomparison of model-based projections for the new IPCC emission scenarios. The Finnish Environment 644. Finnish Environment Institute.

Sengupta, S., & Boyle, J. S., 1998. Using Common Principal Components in Comparing GCM Simula-

tions. Journal of Climate, 11, 816–830.

Torrence, C., & Compo, G.P., 1998. A Practical Guide to Wavelet Analysis. Bull. Amer. Meteor. Soc., 79, 61–78.

Tuomenvirta, H., Drebs, A., Førland, E., Tveito, O.E., Alexandersson, H., Laursen, E.V., & Jónsson, T., 2001. Nordklim data set 1.0. KLIMA 08/01. met.no, P.O.Box 43 Blindern, N-0313 Oslo, Norway (www.met.no).